nature

THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE

A MAGNETIC MOMENT

Individual polarons imaged in a two-dimensional lattice PAGE 358

ASTRONOMY

AFTER THE BIG BANG

The hunt for signals from the very early Universe

PAGE 288

MICROBIOLOGY

STERILE ENVIRONMENT

Human placenta found to be free of bacteria

PAGES 317 & 329

CELL BIOLOGY

HOW TO SENSE SALT

A membrane lipid that helps plants perceive salinity

PAGES 318 & 341

NATURE.COM

15 August 2019

Vol. 572, No. 7769

THIS WEEK

EDITORIALS

WORLD BANK The Pandemic Emergency Financing Facility is not fit for purpose **p.285**

MY NAME IS HERACLES Metre-high flightless parrot lived in New Zealand **p.287**

STRESSED OUT Water stress is a fact of life for more than 2 billion people **p.288**

Time to listen to climate advice

The Intergovernmental Panel on Climate Change has done its job. Now, decision makers must do theirs — and a nascent youth movement is showing them how.

It isn't often that a climate report is this well timed. The Intergovernmental Panel on Climate Change (IPCC) review on climate and land use, released last week (see page 291), has arrived in time for several international meetings on the future of the environment. This August and September, government representatives will gather under the United Nations umbrella in Nairobi, New Delhi and New York City to review progress in protecting biodiversity and mitigating desertification and climate change. The IPCC's latest warnings should turbocharge those deliberations.

Between 2007 and 2016, food production, agriculture, forestry and other human activities related to land use accounted for 21–37% of anthropogenic, or human-caused, greenhouse-gas emissions, the IPCC review says. These emissions could be reduced, it adds, if more land was available to absorb carbon. This could be achievable if more consumers reduced their meat consumption in favour of plant-based diets; more forests were protected and managed sustainably; and soils were replenished with organic content.

But this is as far as the IPCC's authority goes. The panel's job is to describe what humans are doing to the climate. It can suggest how to slow down or reverse these effects, and how humans might adapt to a warming world. The IPCC can make suggestions, but turning these into action is beyond its remit.

When it comes to the role of international political leadership in tackling climate change, the record of achievement leaves much to be desired. But now, because of the IPCC's findings, and with the help of a vigorous youth climate movement — which, unlike adult policymakers, seems to actually pay attention to the IPCC — an opportunity has arisen for real action.

Take the UN Convention on Biological Diversity, representatives of which will gather in Nairobi later this month. A decade ago, the convention's member countries set themselves a 2020 deadline to address the underlying causes of biodiversity loss. Despite the impending deadline, progress has been limited. Delegates will consider extending the deadline and, potentially, setting new targets. But biodiversity is dwindling, in large part, because industrial-scale farming and broader industry is destroying and polluting habitats. As long as these issues remain, an extension is unlikely to make a difference.

At the beginning of next month, it will be the turn of countries belonging to the UN Convention to Combat Desertification (UNCCD) to meet in New Delhi. Desertification happens when land in already-dry parts of the world is degraded through the loss of productive soils. Its human causes include over-cultivation, overgrazing, deforestation and poor irrigation.

The UNCCD's member countries will consider a proposal to integrate their work in combating desertification with the UN's Sustainable Development Goals — a move that should be encouraged. This would avoid duplication of effort, and could speed up progress. But, as the latest IPCC report indicates, droughts in dryland regions have been increasing, on average, by slightly more than 1% per year since 1961. And climate change is making land degradation worse.

Last, but not least, as September draws to a close, world leaders will assemble in New York City for a climate summit convened by UN secretary-general António Guterres, where the IPCC's latest findings will also be considered. As the IPCC report points out, the global mean surface temperature increased by about 0.87 °C (with a likely range of 0.75–0.99 °C) between 1850 and 2015. Guterres wants leaders to come to New York with concrete plans to reduce greenhouse-gas emissions by 45% over the next decade, and to reach net zero by 2050. But whether they are capable of this — or willing to do so — is an open question.

The IPCC can make suggestions, but turning these into action is beyond its remit.

Combating climate change and desertification and slowing the rate of biodiversity loss are even more difficult to achieve, because each respective UN convention is structured to be independent of the others — unlike the reality of threats to biodiversity, climate change and desertification, which are interlinked.

This is where the IPCC's report also stands out. Its authors come from diverse disciplines — and, for the first time, a majority are from developing countries. They have engaged in detailed conversations and produced a document that integrates perspectives on biodiversity and desertification, as well as food and agriculture, into its analysis and findings. The UN conventions could do much more to adopt such an approach.

YOUNG PEOPLE CARE ABOUT CLIMATE

As each of the UN conventions faces continuing challenges, the IPCC can at least be assured of support from the next generation. It has garnered a following among the growing international youth climate movement. Members keenly absorb every new report, including participants in the school strike for climate, led by Swedish teenage activist Greta Thunberg.

Thunberg makes a point of namechecking the IPCC and quoting paragraph and page numbers in speeches, as she did in an address to the French parliament at the end of last month.

As government delegates get ready for Delhi, Nairobi and New York, they must prepare to answer why, if children can understand the meaning of the IPCC assessments, adults cannot do the same?

The youth climate movement's members are brave, and they are right. It has been almost three decades since the three UN conventions — on biodiversity, climate and desertification — were agreed at the Earth Summit in Rio de Janeiro. And it has been 31 years since the IPCC was created to advise decision makers. Yet environmental promises have not been matched by meaningful action.

Younger generations know, perhaps better than the adults, that the world might not have another three decades to prevent climate impacts that will be even more serious than those we face now. Politicians must act now. ■

HARVARD T. H. CHAN SCHOOL OF PUBLIC HEALTH

Pandemic bonds: designed to fail in Ebola

The World Bank's funding scheme for disease outbreaks drained potential resources from the Democratic Republic of the Congo, says **Olga Jonas**.

The final toll of the Ebola outbreak in West Africa in 2014–16 was more than 11,000 lives, plus an estimated US$53 billion from economic disruption and collapse of health systems. In the outbreak's wake, the global health community scrambled to deliver initiatives for increased health security. One flagship programme was the World Bank's Pandemic Emergency Financing Facility (PEF). Under the scheme, investors who buy pandemic bonds receive generous 'coupons', which annually pay about 13% interest. This compensates investors for the risk that the bonds will make 'insurance' payouts to fight pandemics under certain conditions. Otherwise, cash returns to the investors when the bonds mature in July 2020.

The world's second-largest Ebola outbreak, in the Democratic Republic of the Congo (DRC), has now entered its 13th month and has caused at least 1,800 deaths. In July, the World Bank announced that it would, independently of PEF mechanisms, mobilize up to $300 million towards the Ebola outbreak. Meanwhile, the PEF has cost much more than it has brought in. The World Bank, where I worked for 3 decades as an economist, has not advertised the bonds' exact terms, but I have ploughed through the confusing 386-page bond prospectus. The PEF has already paid around $75.5 million to bondholders as premiums, but has not disclosed how much they have been paid in interest — and it is set to pay much more. However, outbreak responders have received just $31 million from the PEF, and the much-touted potential payout of $425 million is highly unlikely. Twice as many investors signed up to buy pandemic bonds as were available. It was a good deal for investors, not for global health. Absurdly, discussions on a second PEF are under way.

IT WAS A GOOD DEAL FOR INVESTORS, NOT FOR GLOBAL HEALTH.

The PEF was backed by about $190 million in donations from 3 countries and the World Bank's International Development Association (IDA), a fund that provides around $20 billion to the world's 75 or so poorest countries each year. All the resources devoted to the PEF would have been better used elsewhere. Instead of spending its funds and attention on partnering with reinsurance firms, the IDA should have focused on improving public-health capacity directly or on building up the Contingency Fund for Emergencies at the World Health Organization (WHO) so that all money would go to countries in need. Former World Bank chief economist and US treasury secretary Larry Summers described the PEF as "financial goofiness" motivated by government and World Bank officials eager to boast about a creative initiative that engaged the private sector.

Early action against outbreaks is imperative because it is both more effective and less costly. But making the bonds attractive to investors meant designing them to reduce the probability of payout. The PEF stipulates a payout of $45 million for Ebola if the officially confirmed death toll reaches 250 (which occurred in the DRC by mid-December last year), but only if at least 20 deaths occurred in a second country. Given that the WHO lists only one multi-country outbreak amid more than 30 that occurred in a single country, this requirement is inappropriate. The DRC is much bigger and more populous than all three countries involved in the West African outbreak.

The World Bank has said that the PEF is working as intended by offering the potential of 'surge' financing. Tragically, current triggers guarantee that payouts will be too little because they kick in only after outbreaks grow large. What's more, fanfare around the PEF might have encouraged complacency that actually increased pandemic risk. Following false assurance that the World Bank had a solution, resources and attention could shift elsewhere.

Rather than a lack of funds, vigilance and public-health capacity have been the main deficiencies. When governments and the World Bank are prepared to respond to infectious-disease threats, money flows within days. In the 2009 H1N1 influenza outbreak in Mexico, clinics could diagnose and report cases of disease to a central authority that both recognized the threat and reacted rapidly. The Mexican government requested $25.6 million from an existing World Bank-financed project for influenza response and received the funds the next day.

For the 2014–16 Ebola outbreak, substantial funds started flowing nine months after it began. Financing was slow because the affected countries, the World Bank and the WHO were not adequately monitoring the disease, and global health leaders did not pay attention until the outbreak became a full-blown crisis.

Increasing surveillance, diagnostics and other capacities for response to outbreaks will do more than flashy financing schemes to reduce threats from infectious disease — including antimicrobial resistance. World Bank analyses show that poor countries' investments in core veterinary and human public-health systems bring returns of 25–88% annually. The World Bank can provide robust financing and operational support for such investment; it should make this a priority.

The Ebola outbreak in West Africa should have been a sufficient wake-up call for the international community to establish a plan to get ahead of outbreaks. There have been important improvements since 2016, including reforms of WHO emergency programmes, and external evaluations of individual countries' core public-health capacities.

But the best investment of funds and attention is in ensuring adequate and stable financing for core public-health capacities. The PEF has failed. It should end early — and IDA funds should go to poor countries, not investors. ■

Olga Jonas *is a senior fellow and economic adviser at the Harvard Global Health Institute in Cambridge, Massachusetts.*
e-mail: olga_jonas@harvard.edu

SEVEN DAYS *The news in brief*

ENVIRONMENT

Running dry

More than one-third of the global population lives in countries under "high" or "extremely high" water stress, according to an analysis by the World Resources Institute (WRI), an environmental think tank in Washington DC. The analysis, released on 6 August, found that 17 countries are extremely water-stressed and each year use more than 80% of their total water supply — from surface and groundwater sources (see go.nature.com/2yxfmxq). The 27 countries under high water stress use 40–80% of their annual supply. The WRI collected more than 50 years of global data on water use and supplies to produce the Aqueduct Water Risk Atlas. The institute found that even countries with low average levels of water stress can have hotspots, with states or regions under extremely high stress. In the United States, which ranks 71st on the WRI's list, New Mexico's level of water stress is on a par with that of Eritrea.

EVENTS

Drug data

Manipulated data have tainted the approval of a gene therapy widely regarded as the most expensive drug in the world. But the US Food and Drug Administration (FDA) said on 6 August that it should remain on the market while the agency assesses the situation. In May, the FDA approved Zolgensma (onasemnogene abeparvovec-xioi) to treat the most severe form of spinal muscular atrophy, a leading genetic cause of infant death. Before the approval, the company that developed the drug, AveXis of Bannockburn, Illinois, had discovered manipulated data in animal studies performed to help establish a production process. AveXis, now owned by Swiss pharmaceutical giant Novartis, did not disclose the finding until 28 June. In a statement, Novartis said that the delay was to allow for an internal investigation, and that the data in question are limited to an older process no longer used to produce the therapy. Treatment with Zolgensma costs more than US$2 million.

JEROME DELAY/AP/SHUTTERSTOCK

Two Ebola drugs show promise in trial

Two Ebola drugs have proved so effective in a clinical trial that researchers will make the treatments available to anyone infected with the virus in the Democratic Republic of the Congo (DRC), where Ebola has killed nearly 1,900 people over the past year. The survival rate for people who received either drug shortly after infection, when levels of the virus in their blood were low, was 90%, officials with the World Health Organization and the US and DRC governments said on 12 August. One of the drugs, REGN-EB3, is a cocktail of three monoclonal antibodies against Ebola made by Regeneron Pharmaceuticals in Tarrytown, New York. The second, mAb114, is derived from a single antibody recovered from the blood of a person who survived Ebola in the DRC in 1995, and was developed by the US National Institute of Allergy and Infectious Diseases. The trial of the drugs, and two others, began last November.

PUBLISHING

Elsevier dispute

At least 30 professors at the University of California (UC) system have stepped down from the editorial boards of Elsevier's flagship journals because of a disagreement over open access. Negotiations to renew the institution's subscriptions to the publishing giant's journals broke down in February owing to a dispute about the cost of making UC-produced research papers freely available. Last month, Elsevier cut off UC academics' access to new papers published in its journals. In a letter on 7 August, prominent UC researchers, including CRISPR pioneer Jennifer Doudna and Nobel prizewinner Elizabeth Blackburn, said that they would no longer provide editorial services for Elsevier's 28 Cell Press journals until a new contract is signed.

POLICY

Science visas

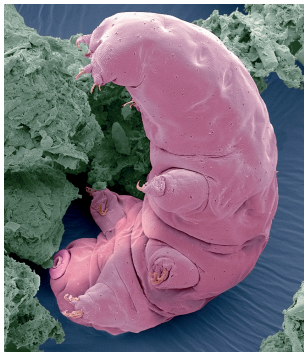
The United Kingdom will develop a new fast-track visa route for scientists, Prime Minister Boris Johnson said on 8 August. The government is exploring measures such as abolishing a cap on the Exceptional Talent visa route, and removing the need for scientists to hold a job offer before arriving in Britain.

(*Nature* found in 2018 that the Exceptional Talent visa route was vastly underused.) Science leaders welcomed the move but stressed that Brexit remained overwhelmingly negative for research. Half of all foreign academic scientists in the United Kingdom are from the European Union and do not require a visa to enter or work. The government also pledged to make up some EU research funding that would be lost if the United Kingdom leaves the EU without a deal on 31 October. It said that it would assess UK funding applications under review by the EU on that date, and fund those deemed successful. The funding could be worth €600 million (US$672 million) to UK scientists, says Universities UK, which represents the country's universities.

SPACE

Moon bears

The Israeli spacecraft Beresheet, which crash-landed on the Moon in April, delivered thousands of millimetre-sized tardigrades (**pictured**) to the lunar surface, it emerged this week. The hardy creatures, also known as water bears, were part of an archive created by the Arch Mission Foundation, a non-profit organization in Los Angeles, California, that



STEVE GSCHMEISSNER/GETTY

aims to preserve a backup of Earth's culture and species. The foundation says that calculations of the likely energy of the crash suggest that the DVD-sized archive, which is stronger than a 'black box' flight recorder, probably survived intact. Ultraviolet radiation on the Moon would kill the tardigrades. But if they remain in the archive or are buried, they might survive in a desiccated state, from which they can later be revived. Unlike Mars, the Moon is thought to be inhospitable to life, and no protections exist to prevent spacecraft from contaminating its surface.

PEOPLE

Science minister

South Korea's President Moon Jae-in has nominated Choi Ki-young, who leads an effort to create semiconductor chips that mimic how the brain works, to be minister of science, information and communication technologies. Choi's nomination comes amid growing tensions between South Korea and Japan. In July, Japan imposed export restrictions on materials such as photoresists and hydrogen fluoride, which are crucial to South Korea's semiconductor and electronic-display industries. Choi said he felt a "heavy responsibility" in assuming the position given the situation with Japan, according to Yonhap News Agency.

FACILITIES

Lab closure

The US Army Medical Research Institute of Infectious Diseases (USAMRIID), which studies dangerous pathogens such as Ebola and plague, has halted operations indefinitely after government inspectors found problems with its wastewater disposal systems and personnel-certification records. On 18 July, inspectors from the US Centers for Disease Control and Prevention (CDC) sent a letter to the laboratory in Fort Detrick, Maryland, ordering it to immediately suspend all research with dangerous pathogens and toxins. The closure was first reported by Maryland local newspaper *The Frederick News-Post* on 2 August. A USAMRIID spokesperson said that no infectious agents had been detected outside the containment areas, and that the facility is "continuing to work closely with the CDC on corrective actions".

FUNDING

UK university grants

An inquiry by the UK House of Lords has found that funding for research at universities is under threat. In a report published on 8 August, the Lords Science and Technology Committee told the UK government that the 'block grant' for universities — money awarded to institutions for research on the basis of the quality of their work — has fallen by 13% in real terms since 2010. It added that a recommendation in May to cut university tuition fees would have severe financial consequences for science, because other income streams that support research would be diverted to make up for the shortfall in teaching funds. The committee urged the government to address the deficit in research funding.

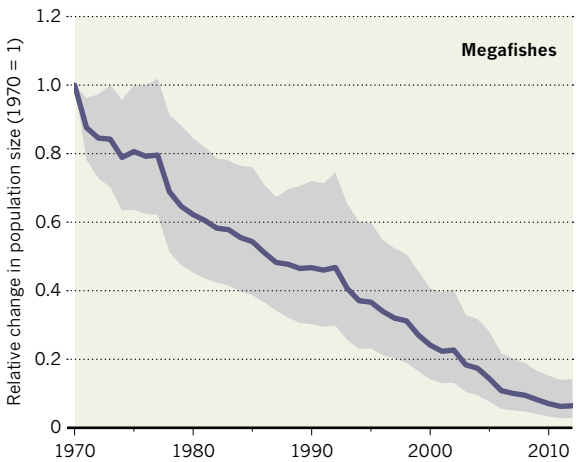
TREND WATCH

Freshwater megafishes — giants weighing more than 30 kilograms that can live for decades — declined by more than 94% between 1970 and 2012, according to a study. The findings, published on 8 August in the journal *Global Change Biology*, are part of an analysis that looked at the populations of enormous freshwater animals in the world's rivers and lakes. The drop-off reflects a broader downward trend in the populations of freshwater megafauna — such as caimans, giant salamanders and giant catfish — around the world.

The study authors collected data on the populations of 126 large freshwater species from 72 countries, and estimate that the populations of big freshwater animals have fallen by 88%. They expected megafishes to be hit the hardest by human activities such as overfishing and loss of habitat, because many giant fish species mature late, have relatively few offspring and require large, intact habitats for migration. Their movements are increasingly hampered by hydroelectric dams in the world's greatest river basins, such as the Mekong, Congo, Amazon and Ganges.

PLUNGING POPULATIONS

Habitat loss and hunting led to steep population declines in huge freshwater species, such as the Mekong giant catfish (*Pangasianodon gigas*), between 1970 and 2012.



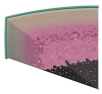
SOURCE: F. HE *ET AL. GLOB. CHANGE BIOL.* HTTPS://DOI.ORG/10.1111/GCB.14753 (2019).

NEWS IN FOCUS

ASTRONOMY What's next for the beleaguered Thirty Meter Telescope? **p.292**

ORIGINS Microbe grown in lab could help to explain rise of complex life **p.294**

FUNDING Austerity in Mexico pushes science to breaking point **p.294**



ASTRONOMY The first billion years of the Universe begin to yield secrets **p.298**

YASUYOSHI CHIBA/AFP/GETTY

Cattle farming, shown here in northern Brazil, is emission-intensive and often accompanies large-scale deforestation.

CLIMATE CHANGE

Eat less meat: UN climate-change panel tackles diets

Report on climate change and land comes amid accelerating deforestation in the Amazon.

BY QUIRIN SCHIERMEIER

Efforts to curb greenhouse-gas emissions and the impacts of global warming will fall significantly short without drastic changes in global land use, agriculture and human diets, researchers warn in a high-level report commissioned by the United Nations.

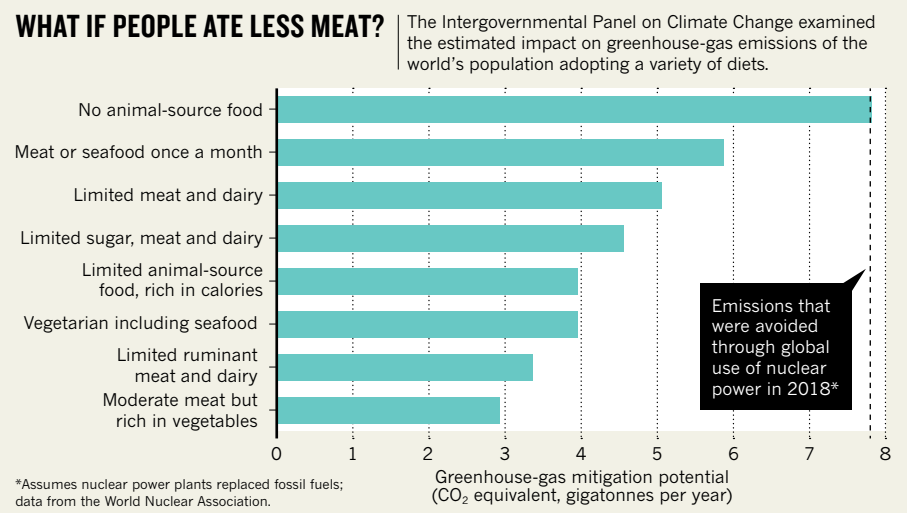
The special report by the Intergovernmental Panel on Climate Change (IPCC) describes plant-based diets as a major opportunity for mitigating and adapting to climate change — and includes a policy recommendation to reduce meat consumption.

On 8 August, the IPCC released a summary of the report, which is designed to inform upcoming climate negotiations amid the worsening global climate crisis. More than 100 experts, around half of whom hail from developing countries, worked on the report.

"We don't want to tell people what to eat," says Hans-Otto Pörtner, an ecologist who co-chairs the IPCC's working group on impacts, adaptation and vulnerability. "But it would indeed be beneficial, for both climate and human health, if people in many rich countries consumed less meat, and if politics would create appropriate incentives to that effect."

Researchers also note the relevance of the report to tropical rainforests. The Amazon rainforest is a huge carbon sink that acts to cool global temperature, but rates of deforestation are accelerating, in part because of the policies and actions of the government of Brazilian President Jair Bolsonaro.

Unless stopped, deforestation could turn much of the remaining Amazon forests into a degraded type of desert, and could release more than 50 billion tonnes of carbon into the atmosphere in 30 to 50 years, says Carlos Nobre, a climate scientist at the University of São Paolo in Brazil. "That's very worrying." ▶



"Unfortunately, some countries don't seem to understand the dire need of stopping deforestation in the tropics," says Pörtner. "We cannot force any government to interfere. But we hope that our report will sufficiently influence public opinion to that effect."

Although the burning of fossil fuels garners the most attention, activities relating to land management produce almost one-quarter of heat-trapping gases resulting from human activities. The race to limit global warming to 1.5 °C above pre-industrial levels — the goal of the international Paris climate agreement made in 2015 — might be a lost cause unless land is used in a more climate-friendly way, the latest IPCC report says.

Cattle are often raised on pastures created by clearing woodland, and produce methane, a potent greenhouse gas, as they digest their food. The report states with high confidence that balanced diets featuring plant-based and sustainably produced animal-sourced food "present major opportunities for adaptation and mitigation while generating significant co-benefits in terms of human health".

By 2050, dietary changes could free up several million square kilometres of land, and reduce global carbon dioxide emissions by up to eight billion tonnes per year, relative to business as usual, the scientists estimate (see 'What if people ate less meat?').

"It's really exciting that the IPCC is getting such a strong message across," says Ruth Richardson in Toronto, Canada, who is the executive director at the Global Alliance for the Future of Food, a coalition of philanthropic foundations.

The report cautions that land must remain productive to feed a growing world population. Warming enhances plant growth in some regions, but in others — including northern Eurasia, parts of North America, Central Asia and tropical Africa — increasing water stress seems to reduce vegetation. So the use of biofuel crops and the creation of new forests — measures that could mitigate global warming — must be carefully managed to avoid food shortages and biodiversity loss, the report says.

FLOODS AND DROUGHTS

Farmers and communities around the world must also grapple with more-intense rainfall, floods and droughts resulting from climate change, warns the IPCC. Land degradation and expanding deserts threaten to affect food security, increase poverty and drive migration.

About one-quarter of Earth's ice-free land area seems to be suffering from human-induced soil degradation already — and climate change is expected to make things worse.

The report might provide a much-needed, authoritative call to action, says André Laperrière, the executive director of Global Open Data for Agriculture and Nutrition in Wallingford, UK. Nobre hopes that the IPCC's voice will give greater prominence to land-use issues in upcoming climate talks. "I think that the policy implications of the report will be positive in terms of pushing all tropical countries to aim at reducing deforestation rates," he says.

Governments from around the world will consider the IPCC's findings at a UN climate summit next month in New York City. The next round of climate talks of parties to the Paris agreement will take place in December in Santiago. "We need to mainstream climate-change risks across all decisions," said António Guterres, the UN secretary-general. "That is why I am telling leaders don't come to the summit with beautiful speeches." ■

ASTRONOMY

What's next for the embattled Thirty Meter Telescope?

Protesters on Hawaii's Big Island have prevented construction for a month.

BY ALEXANDRA WITZE

A stand-off over plans to build a mega-telescope on Hawaii's tallest mountain has entered its fifth week and shows no signs of stopping. Hundreds of protesters are blocking access to Mauna Kea, the mountain on Hawaii's Big Island where construction of the Thirty Meter Telescope (TMT) was set to begin on 15 July.

The US$1.4-billion telescope's enormous light-gathering mirror — nine times the area of those in today's biggest telescopes — will allow it to peer at stars and galaxies with unprecedented sharpness. That will allow scientists to explore fundamental questions such as how galaxies arose in the early Universe and what planets around distant stars look like.

Here, *Nature* examines how the fight over the telescope could evolve.

Who are the protesters, and what do they want?

The activists who oppose the TMT encompass a broad swathe of the Hawaiian community, including university professors, local leaders and students. Most are Native Hawaiians. Their protests have garnered widespread support from people in and beyond Hawaii, including celebrities of Asian–Pacific ancestry such as actor Jason Momoa, who visited the encampment on 31 July.

The protesters do not want the TMT to be built on Mauna Kea. They say they are protecting the site, which is sacred to Native Hawaiians and already hosts 13 observatories (5 of which are supposed to be dismantled

Protesters in Hawaii have blocked access to the mountain of Mauna Kea.

RONIT FAHL/ZUMA

before the TMT begins operations).

"We have always been here and we will always be here," said Kealoha Pisciotta, a protest leader, during a press conference on 18 July. "The TMT will never be built."

Many other Native Hawaiians do support the project. And a poll of 1,367 state residents, released on 7 August by the *Honolulu Civil Beat* newspaper, found that 64% supported the project, whereas 31% opposed it.

Hasn't this been going on for a while?

Months-long protests in 2015 scuttled the TMT project's first attempt to build on Mauna Kea. In 2018, after further legal challenges to the TMT's right to proceed, Hawaii's supreme court ruled that the telescope's construction permit was valid. That move set the stage for the attempt last month to start construction.

The current stand-off has been more intense than the 2015 protests in two important ways: it has drawn more activists to the mountain, and it shut down activity at the telescopes already on Mauna Kea for more than three weeks.

How have scientists reacted?

Many scientists have spoken out against building the TMT in Hawaii, citing the need to listen to indigenous voices. They include a number of students and researchers affiliated with institutions working on the TMT. The president of the University of British Columbia in Vancouver, which is participating in the TMT project as a member of the Association of Canadian Universities for Research in Astronomy, has called for a 60-day moratorium on the project.

Other researchers, including two officials with the Canadian Astronomical Society, say that the TMT project should work towards building in Hawaii. The project should pursue a site on Mauna Kea "for as long as there remains a realistic possibility to peacefully negotiate a route for this to happen, and to do so in a way that means the project is broadly welcomed and viable in Hawaii", astronomers Michael Balogh at the University of Waterloo in Ontario and Rob Thacker at Saint Mary's University in Halifax, both in Canada, wrote to society members on 1 August.

TMT officials say they are hopeful that the project can move forwards.

"We've been through a ten-year process, and it's urgent for us to get started," says Gordon Squires, vice-president of external affairs for the Thirty Meter Telescope International Observatory, the formal name for the telescope project. "We have a lot of respect for everybody — those who oppose us and those who support us — and are looking forward to a safe resolution to this."

What about the telescopes that are already on Mauna Kea?

They were shuttered on 16 July, the second day of protests, when it became clear that workers would not be able to regularly go up and down the mountain. On 9 August, observatory leaders announced that they had reached an agreement with the activists to allow limited operations to resume. The telescopes are slowly coming back online, and it could be weeks before they are back to observing as normal.

The interruption to scientific activity on Mauna Kea was the longest in the five decades of astronomy on the mountain.

How have officials in Hawaii responded?

David Ige, Hawaii's governor, issued an emergency proclamation on 17 July that gave police greater power to restrict access to Mauna Kea and deploy additional officers, among other things. On that day, law-enforcement officials arrested and released 38 protesters, most of them Native Hawaiian elders.

On 30 July, Ige rescinded the proclamation, saying that conditions on the mountain had changed and it was no longer necessary. He also extended the window in which the TMT's construction could start by two years, to September 2021. That gives the project more time to negotiate a solution to the impasse.

Ige has put Harry Kim, the mayor of Hawaii County, in charge of figuring out what to do next. Kim has been holding meetings with a broad swathe of community leaders to discuss possible future steps.

Can the TMT be built somewhere else?

The project does have a backup site: the Roque de los Muchachos Observatory on La Palma, one of Spain's Canary Islands. The community in La Palma has mostly been supportive, and Spain's minister of science, the former astronaut Pedro Duque, said last month that the TMT is welcome there. But the environmental group Ecologists in Action has been speaking out against the idea of building the telescope on La Palma, saying that it would harm a natural area of great value.

There are some drawbacks to the La Palma site. Because it is lower in elevation than Mauna Kea — 2,250 metres as opposed to 4,050 metres — the TMT would need to peer through more of Earth's atmosphere. Having more water vapour between the TMT and the stars would reduce the quality of the telescope's observations.

And the TMT project has not yet finalized all the agreements with the local government that would allow construction of the telescope on La Palma. On 5 August, TMT executive director Ed Stone confirmed that the project has applied for a building permit at La Palma, to help keep that option open.

What would need to happen for the project to relocate there?

The TMT board, which includes representatives from two California universities and the governments of Canada, China, India and Japan, would need to approve the move.

One complicating factor is that the project will probably need hundreds of millions of dollars from the US National Science Foundation to finish its construction. US legislators might be less willing to fund the TMT if it is not built on US soil. For Japan, China and India, the Canary Islands site is farther away and less desirable than Hawaii. ■

MICROBIOLOGY

Cells hint at roots of complex life

'Asgards' isolated and grown in the lab could be similar to cells that evolved into eukaryotes.

BY JONATHAN LAMBERT

Scientists in Japan report that they have isolated and grown microbes from an ancient lineage of archaea — single-celled microbes that look, superficially, like bacteria but are quite distinct — that was previously known only from genomic sequences.

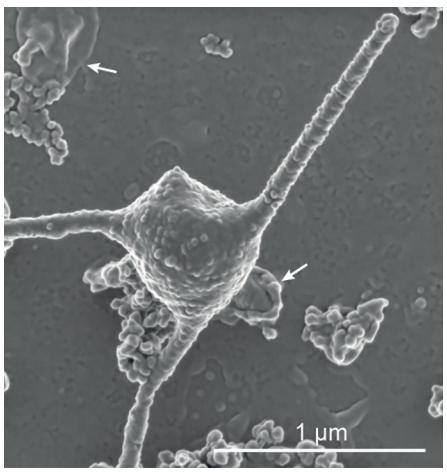
The work, posted online as a preprint (H. Imachi *et al.* Preprint at bioRxiv http://doi.org/gf5z2n; 2019), gives scientists their first look at the kinds of organism that could have made the jump from simple, bacteria-like cells to eukaryotes — the group of organisms whose cells have nuclei and other structures, and which includes plants, fungi and humans.

"This is a monumental paper that reflects a tremendous amount of work and perseverance," says Thijs Ettema, an evolutionary microbiologist at Wageningen University in the Netherlands.

The mysterious group, called Lokiarchaea, rose to prominence from microbial muck dredged up off the coast of Greenland. In 2015, Ettema and his colleagues sequenced genetic fragments from the sediment and assembled them into fuller genomes of individual species (A. Spang *et al. Nature* **521**, 173–179; 2015).

One genome was clearly a member of the archaea, but also had some eukaryote-like genes. The researchers called it Lokiarchaea, after Loki, the trickster of Norse mythology.

Soon, other labs found more Loki-like archaea, and together these formed the Asgard archaea, named after a mythological region inhabited by Norse gods. Many analyses suggest that some distant Asgard-like ancestor gave rise to all eukaryotes.



H. IMACHI *ET AL.*

Wisp-like protrusions make this candidate new strain look like 'an organism from outer space'.

Proponents of this view think that, some 2 billion years ago, an Asgard-like archaeon gobbled up a bacterium, sparking a mutually beneficial relationship known as endosymbiosis. The bacterium would have evolved into mitochondria, the 'powerhouse' organelles of the cell that helped to fuel eukaryotes' rise.

But no one had succeeded in growing Asgards in the lab.

To cultivate sea-floor microbes, Hiroyuki Imachi, a microbiologist at the Japan Agency for Marine-Earth Science and Technology in Yokosuka and his collaborators built a bioreactor that mimicked the conditions of a deep-sea methane vent. They then waited 5 years for the slow-growing microbes to multiply.

Next, they placed samples from the reactor, along with nutrients, in glass tubes, which sat for another year before showing any signs of life. Genetic analysis revealed a barely perceptible population of Lokiarchaea. The researchers patiently coaxed the Lokiarchaea into higher abundance and purified the samples.

Finally, after 12 years of work, the scientists produced a stable lab culture containing only this new Lokiarchaeon and a different methane-producing archaeon. The authors declined requests for interviews from *Nature*'s news team while their paper was under review at a journal.

Like other archaea and bacteria, Asgards have relatively simple interiors, but their external surface can produce wisp-like protrusions. "I don't think anyone predicted that it would look like this," says Ettema. "It's sort of an organism from outer space."

The team reports that the cultured Lokiarchaeon produces energy by breaking down amino acids, as predicted from genomic studies. And, because the researchers could extract and sequence DNA from a pure sample, rather than sediment containing a multitude of organisms, their findings could confirm that Lokiarchaea do contain numerous eukaryote-like genes.

Ettema says that many more Asgards will need to be cultured for researchers to work out whether, and how, Asgard-like archaea gave rise to eukaryotes. ■

FUNDING

Mexican science suffers under budget cuts

Research institutes are rationing electricity to save money.

BY GIORGIA GUGLIELMI

Austerity measures recently enacted by Mexico's president are pushing the country's scientific efforts — chronically underfunded for years — to a breaking point, according to researchers.

As part of broader cost-cutting measures aimed at freeing up money for poverty-alleviation programmes, in May, President Andrés Manuel López Obrador cut 30–50% of the money that federally funded institutions — including centres supported by Mexico's main research funding agency, the National Council of Science and Technology (CONACYT) — spend on travel, petrol, office supplies and salaries for temporary workers.

Several research institutes say that, since then, they have rationed electricity and sacked temporary workers. Scientists have cancelled conference travel and international projects, and have relied on crowdfunding campaigns to pay for supplies. The monetary uncertainty has also deterred Mexican researchers working abroad from returning to take jobs at home.

The measures came on top of a roughly 12% cut to the 2019 budget for CONACYT that López Obrador's administration enacted in December 2018. The move left the agency with 18.8 billion pesos (US$960 million).

"Mexican science has never been well funded," says Antonio Lazcano, a biologist at the National Autonomous University of Mexico (UNAM) in Mexico City. But the austerity measures, on top of the cuts to CONACYT's budget, threaten to hamper the recruitment of early-career researchers, as well as the monitoring efforts for potential disasters such as

earthquakes and epidemics, he says. Without advances in science and technology — which drive innovation and attract investors — the cuts could also set back economic growth in Mexico, he adds.

In June, Lazcano and 56 other Mexican scientists wrote an open letter to the government urging officials to reverse these recent funding cuts. As of 13 August, more than 19,000 people had signed the letter online.

RIPPLE EFFECTS

Juan Martínez, an ecologist at the Institute of Ecology in Xalapa, says that the cuts enacted in May are pushing the institute to its limit. "We don't have money to pay [for] electricity," says Martínez, who has signed the open letter. To save energy, the institute has banned employees from charging their phones, turning on the air conditioning, working past 6 p.m. during the week or coming in over the weekend.

Cuauhtémoc Sáenz-Romero, a forest geneticist at the Michoacan University of Saint Nicholas of Hidalgo, worries that he'll have to end collaborations with scientists abroad. He is part of a working group at the Food and Agriculture Organization of the United Nations that is developing improved forest conservation and management strategies across the United States, Canada and Mexico.

The Mexican National Forest Commission was supposed to pay for Sáenz-Romero and two of his colleagues to attend the group's next meeting in Idaho in October. But the commission won't be able to fund the trip. Because the Mexican delegates cannot attend, the meeting has now been cancelled, and it is unclear when it will be rescheduled.

Despite these reports, CONACYT director Elena Álvarez-Buylla insists that the cuts enacted in May are aimed at reducing overspending and will not affect research projects at institutions funded by the agency.

CONACYT plans to have allocated at least 1.6 billion pesos to basic-science projects by the end of 2019, Álvarez-Buylla says. Decisions on new grants will be made at the end of the year, which means that researchers won't get funds until 2020.

Lack of sufficient federal funding in Mexico pre-dates the current administration. Soledad Funes, a molecular biologist at UNAM, says that, over the past decade, calls for basic-science grant applications from CONACYT have been irregular. Funes is currently relying on a 250,000-peso grant provided by her university to continue her research.

Scientists at institutions that don't provide such grants have turned elsewhere for money. Enrique Espinosa, an immunologist at the National Institute for Respiratory Diseases in Mexico City, has started a crowdfunding campaign for money to buy reagents, attend scientific conferences and support a graduate student until they receive a scholarship.

The mounting funding uncertainty has also discouraged Mexican researchers abroad from returning. Jorge Zavala, an astronomer at the University of Texas in Austin, rejected a well-paid academic position at the Institute of Astrophysics, Optics and Electronics in Tonantzintla last year because he wasn't sure how long the money would last.

The post was part of a CONACYT programme covering salaries for young scientists working at Mexican institutions that couldn't afford to pay their researchers. But Zavala wasn't sure whether the programme would have continued under López Obrador's administration.

Zavala plans to apply for academic positions in Europe or the United States in the near future. At some point, he says, "I might go back to Mexico, if things get better." ■

COSMOLOGY

Sky map to plot dark energy

A telescope in Arizona will survey galaxies to reconstruct 11 billion years of cosmic history.

BY DAVIDE CASTELVECCHI

Astronomers are about to embark on their most ambitious galaxy-mapping project ever. Over the next five years, they will use a telescope in Arizona — retrofitted with thousands of small robotic arms — to capture light spectra from 35 million galaxies and reconstruct the Universe's history of expansion. Their main aim: to elucidate the nature of dark energy, the enigmatic force that is pushing the Universe to accelerate at an ever-faster pace.

The Dark Energy Spectroscopic Instrument (DESI) is scheduled to see 'first light' in September. After a commissioning period, its survey of the northern sky — using the 4-metre Mayall Telescope at Kitt Peak National Observatory near Tucson — could start by January 2020. Roughly three-quarters of DESI's US$75-million budget comes from US Department of Energy (DOE), with major contributions from the United Kingdom and France.

DESI is the first in a new generation of experiments investigating the past expansion of the Universe, which come two decades after the first strong evidence of dark energy was found in 1998. Others include ground-based and space observatories set to come online in the 2020s.

The survey will reconstruct 11 billion years of cosmic history. It could answer the first and most basic question about dark energy: is it a uniform force across space and time, or has its strength evolved over eons? ▶

JOHN BRILEY FOR THE WASHINGTON POST/GETTY

The 4-metre Mayall Telescope at Kitt Peak National Observatory near Tucson.

The survey will track cosmic expansion by measuring features of the early Universe, known as baryon acoustic oscillations (BAOs). These oscillations are ripples in the density of matter that left a spherical imprint in space around which galaxies clustered. The distribution of galaxies is highest in the centre of the imprint, a region called a supercluster, and around its edges — with giant voids between these areas.

Superclusters formed in regions where dark matter — invisible material that drives the formation of such large structures — had concentrated under its own gravitational pull.

COSMIC RULER

This primordial pattern of galaxy clustering has remained unchanged since about one million years after the Big Bang. As the Universe grew, BAOs have tracked its expansion; they are now about 320 megaparsecs wide (1 billion light years). Cosmologists use this distance as a ruler; by tracking the size of the BAOs across time, they can reconstruct how the Universe itself expanded.

"The pattern in the map is basically constant; the scale is increasing," says Daniel Eisenstein, a physicist at Harvard University in Cambridge and a spokesperson for DESI.

Tracking BAOs requires a 3D map of galaxies made by measuring their redshifts — the lengthening of the electromagnetic waves in their spectra of light. Redshifts measure how fast a galaxy is receding from the Milky Way, which indicates how far away that galaxy is.

The more redshifts that are measured, the more precise the BAO tracking. Eisenstein and others have found the unmistakable BAO signature in previous galaxy surveys, in particular the US-based Baryon Oscillation Spectroscopic Survey (BOSS) and the Australia-based Two-degree-Field Galaxy Redshift Survey. Together, those surveys mapped nearly 2.4 million galaxies.

"Within a few months, we will surpass what we had for BOSS."

The number of galaxies that DESI will track will eclipse the previous surveys by an order of magnitude. "Within a few months, we will surpass what we had for BOSS," says Michael Levi, a physicist at the Lawrence Berkeley National Laboratory (LBNL) in California and DESI's director.

DESI will achieve such a speed-up thanks to a radically different design. Surveys such as BOSS used optical fibres, placed into holes drilled into custom metal plates, to capture each galaxy's light and deliver it to a separate spectrograph to measure the redshift. But the plates needed to be changed to measure each different part of the sky, which was slow.

DESI will replace the metal plates with 5,000 tiny robotic arms, arranged in a closely packed beehive pattern. Once images of galaxies are projected on the telescope's focal plane — each about 100 micrometres wide — the robotic arms will quickly position optical fibres to within 10 micrometres of the centre of each image, explains Joseph Silber, a mechanical engineer at the LBNL who led the design and construction of the robotic system.

Although BOSS typically changed about five plates a night, DESI's focal plane can be refigured for another part of the sky in a few minutes; the main limitation is how long the exposures need to be to get enough light. Depending on the season and the weather, DESI could take 30 or more exposures, each with thousands of redshifts, in a night.

Other astronomy experiments have used robotic positioners before. But, Silber says, "DESI is definitely the biggest one tried so far."

In addition to probing dark energy, DESI will study dark matter's role in the growth of galaxies and clusters of galaxies by measuring motion in clusters, says DESI spokesperson Nathalie Palanque-Delabrouille, a cosmologist at the French Alternative Energies and Atomic Energy Commission (CEA) Saclay Research Centre outside Paris. This will provide "exquisite tests" of the favourite models of how dark matter drives the growth of large structures, she says. ■

INTO THE DARK AGES

Radioastronomers take aim at the Universe's first billion years.

BY DAVIDE CASTELVECCHI

To get an idea of what the Universe looks like from Earth's perspective, picture a big watermelon. Our Galaxy, the Milky Way, is one of the seeds, at the centre of the fruit. The space around it, the pink flesh, is sprinkled with countless other seeds. Those are also galaxies that we — living inside that central seed — can observe through our telescopes.

Because light travels at a finite speed, we see other galaxies as they were in the past. The seeds farthest from the centre of the watermelon are the earliest galaxies seen so far, dating back to a time when the Universe was just one-thirtieth of its current age of 13.8 billion years. Beyond those, at the thin, green outer layer of the watermelon skin, lies something primeval from before the time of stars. This layer represents the Universe when it was a mere 380,000 years old, and still a warm, glowing soup of subatomic particles. We know about that period because its light still ripples through space — although it has stretched so much over the eons that it now exists as a faint glow of microwave radiation.

The most mysterious part of the observable

JOHN GOLDSMITH/CELESTIAL VISIONS

A night view of part of the Murchison Widefield Array in Western Australia.

Universe is another layer of the watermelon, the section between the green shell and the pink flesh. This represents the first billion years of the Universe's history (see 'An Earth's-eye view of the early Universe'). Astronomers have seen very little of this period, except for a few, exceedingly bright galaxies and other objects.

Yet this was the time when the Universe underwent its most dramatic changes. We know the end product of that transition — we are here, after all — but not how it happened. How and when did the first stars form, and what did they look like? What part did black holes play in shaping galaxies? And what is the nature of dark matter, which vastly outweighs ordinary matter and is thought to have shaped much of the Universe's evolution?

An army of radioastronomy projects small and large is now trying to chart this terra incognita. Astronomers have one simple source of information — a single, isolated wavelength emitted and absorbed by atomic hydrogen, the element that made up almost all ordinary matter after the Big Bang. The effort to detect this subtle signal — a line in the spectrum of hydrogen with a wavelength of 21 centimetres — is driving astronomers to deploy ever-more-sensitive observatories in some of the world's most remote places, including an isolated raft on a lake on the Tibetan Plateau and an island in the Canadian Arctic.

Last year, the Experiment to Detect the Global Epoch of Reionization Signature (EDGES), a disarmingly simple antenna in the Australian outback, might have seen the first hint of the presence of primordial hydrogen around the earliest stars[1]. Other experiments are now on the brink of reaching the sensitivity that's required to start mapping the primordial hydrogen — and therefore the early Universe — in 3D. This is now the "last frontier of cosmology", says theoretical astrophysicist Avi Loeb at the Harvard-Smithsonian Center for Astrophysics (CfA) in Cambridge, Massachusetts. It holds the key to revealing how an undistinguished, uniform mass of particles evolved into stars, galaxies and planets. "This is part of our genesis story — our roots," says Loeb.

A FINE LINE

Some 380,000 years after the Big Bang, the Universe had expanded and cooled enough for its broth of mostly protons and electrons to combine into atoms. Hydrogen dominated ordinary matter at the time, but it neither emits nor absorbs photons across the vast majority of the electromagnetic spectrum. As a result, it is largely invisible.

But hydrogen's single electron offers an exception. When the electron switches between two orientations, it releases or absorbs a photon. The two states have almost identical energies, so the difference that the photon makes up is quite small. As a result, the photon has a relatively low electromagnetic frequency and so a rather long wavelength, of slightly more than 21 cm.

It was this hydrogen signature that, in the 1950s, revealed the Milky Way's spiral structure. By the late 1960s, Soviet cosmologist Rashid Sunyaev, now at the Max Planck Institute for Astrophysics in Garching, Germany, was among the first researchers to realize that the line could also be used to study the primordial cosmos. Stretched, or redshifted, by the Universe's expansion, those 21-cm photons would today have wavelengths ranging roughly between 1.5 and 20 metres — corresponding to 15–200 megahertz (MHz).

Sunyaev and his mentor, the late Yakov Zeldovich, thought of using the primordial hydrogen signal to test some early theories for how galaxies formed[2]. But, he tells *Nature*, "When I went to radioastronomers with this, they said, 'Rashid, you are crazy! We will never be able to observe this'."

The problem was that the hydrogen line, redshifted deeper into the radio spectrum, would be so weak that it seemed impossible to isolate from the cacophony of radio-frequency signals emanating from the Milky Way and from human activity, including FM radio stations and cars' spark plugs.

The idea of mapping the early Universe with 21-cm photons received only sporadic attention for three decades, but technological advancements in the past few years have made the technique look more tractable. The basics of radio detection remain the same;

"THIS IS PART OF OUR GENESIS STORY."

many radio telescopes are constructed from simple materials, such as plastic pipes and wire mesh. But the signal-processing capabilities of the telescopes have become much more advanced. Consumer-electronics components that were originally developed for gaming and mobile phones now allow observatories to crunch enormous amounts of data with relatively little investment. Meanwhile, theoretical cosmologists have been making a more detailed and compelling case for the promise of 21-cm cosmology.

DARKNESS AND DAWN

Right after atomic hydrogen formed in the aftermath of the Big Bang, the only light in the cosmos was that which reaches Earth today as faint, long-wavelength radiation coming from all directions — a signal known as the cosmic microwave background (CMB). Some 14 billion years ago, this afterglow of the Big Bang would have looked uniformly orange to human eyes. Then the sky would have reddened, before slowly dimming into pitch darkness; there was simply nothing else there to produce visible light, as the wavelengths of the background radiation continued to stretch through the infrared spectrum and beyond. Cosmologists call this period the dark ages.

Over time, theorists reckon that the evolving Universe would have left three distinct imprints on the hydrogen that filled space. The first event would have begun some 5 million years after the Big Bang, when the hydrogen became cool enough to absorb more of the background radiation than it emitted. Evidence of this period should be detectable today in the CMB spectrum as a dip in intensity at a certain wavelength, a feature that has been dubbed the dark-ages trough.

A second change arose some 200 million years later, after matter had clumped together enough to create the first stars and galaxies. This 'cosmic dawn' released ultraviolet radiation into intergalactic space, which made the hydrogen there more receptive to absorbing 21-cm photons. As a result, astronomers expect to see a second dip, or trough, in the CMB spectrum at a different, shorter wavelength; this is the signature that EDGES seems to have detected[1].

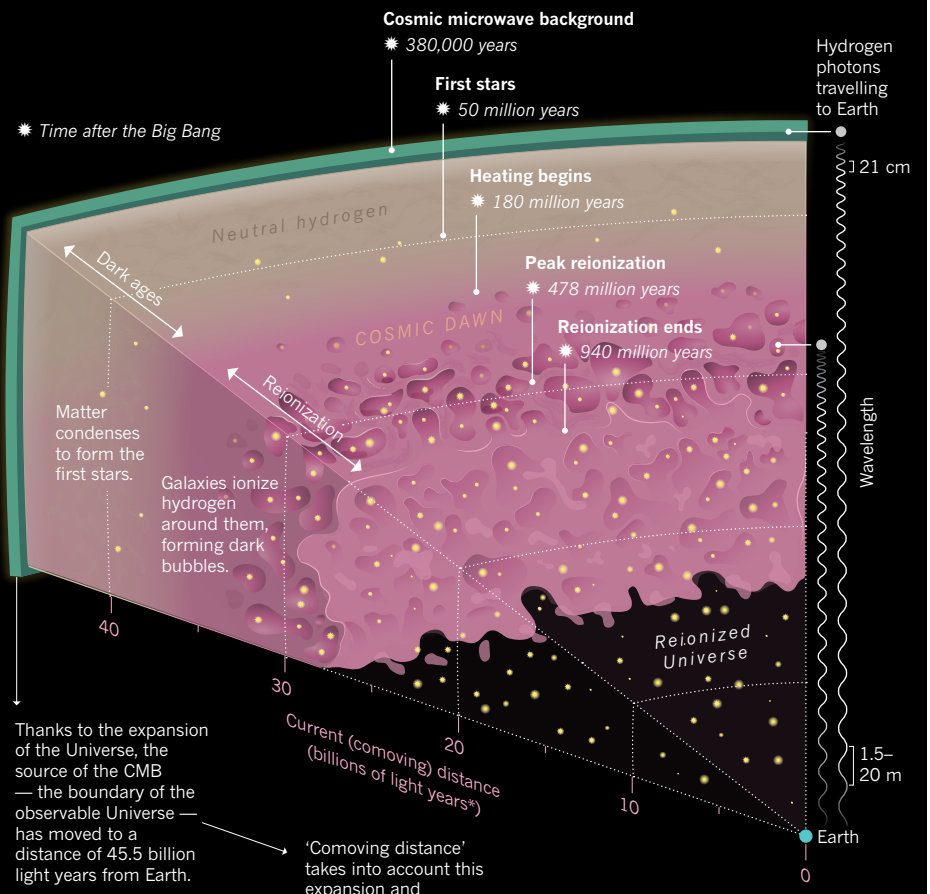
Half a billion years into the Universe's existence, hydrogen would have gone through an even more dramatic change. The ultraviolet radiation from stars and galaxies would have brightened enough to cause the Universe's hydrogen to fluoresce, turning it into a glowing source of 21-cm photons. But the hydrogen closest to those early galaxies absorbed so much energy that it lost its electrons and went dark. Those dark, ionized bubbles grew bigger over roughly half a billion years, as galaxies grew and merged, leaving less and less luminous hydrogen between them. Even today, the vast majority of the Universe's hydrogen remains ionized. Cosmologists call this transition the epoch of reionization, or EOR.

The EOR is the period that many 21-cm radioastronomy experiments, either ongoing or in preparation, are aiming to detect. The hope is to map it in 3D as it evolved over time, by taking snapshots of the sky at different wavelengths, or redshifts. "We'll be able to build up a whole movie," says Emma Chapman, an astrophysicist at Imperial College London. Details of when the bubbles formed, their shapes and how fast they grew will reveal how galaxies formed and what kind of light they produced. If stars did most of the reionization, the bubbles will have neat, regular shapes, Chapman says. But "if there are a lot of black holes, they start to get larger and more free-form, or wispy", she says, because radiation in the jets that shoot out from black holes is more energetic and penetrating than that from stars.

The EOR will also provide an unprecedented test for the current best model of cosmic evolution. Although there is plenty of evidence for

AN EARTH'S-EYE VIEW OF THE EARLY UNIVERSE

The deeper astronomers look into the night sky, the further back in time they see. The oldest observable light is the cosmic microwave background (CMB) — radiation left over from the Big Bang that was emitted when the Universe was just 380,000 years old. Atomic hydrogen formed at that time, and researchers can follow its activities in the early Universe by looking for signs of the radiation that it emitted or absorbed. Hydrogen does this at a characteristic 21-centimetre wavelength, and that radiation has stretched over time as the Universe has expanded. Evidence of that 21-cm signal charts the evolution of the Universe from the dark ages, before the first stars emerged, through to the galaxy-studded cosmos we see today.



Thanks to the expansion of the Universe, the source of the CMB — the boundary of the observable Universe — has moved to a distance of 45.5 billion light years from Earth.

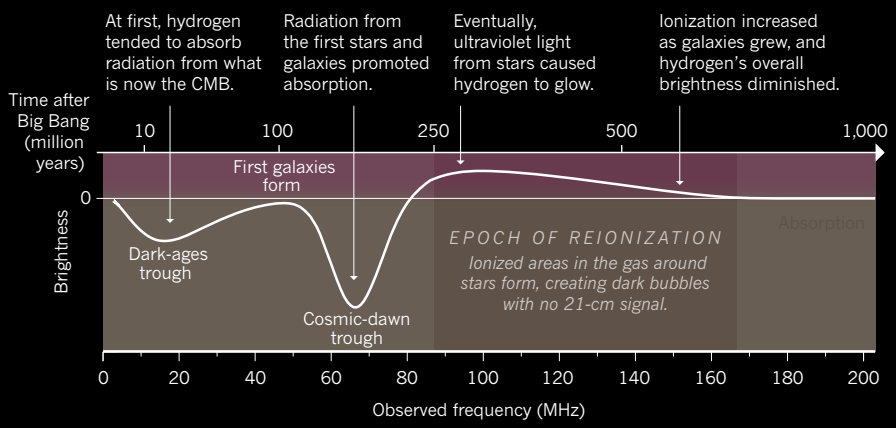
'Comoving distance' takes into account this expansion and represents the distance light has traversed from objects that disappeared long ago.

Although the early cosmos is long gone, its light is only now reaching Earth. The first billion years of cosmic history, still largely unstudied, represent a good 80% of the Universe's observable volume.

* 1 light year = 0.3 parsecs.

IMPRINT OF AN ATOM

This curve represents the overall brightness of hydrogen's 21-cm signal during the first billion years of the Universe's history.



NIK SPENCER/*NATURE*; GRAPH ADAPTED FROM J. R. PRITCHARD & A. LOEB *PHYS. REV. D* **82**, 023006 (2010).

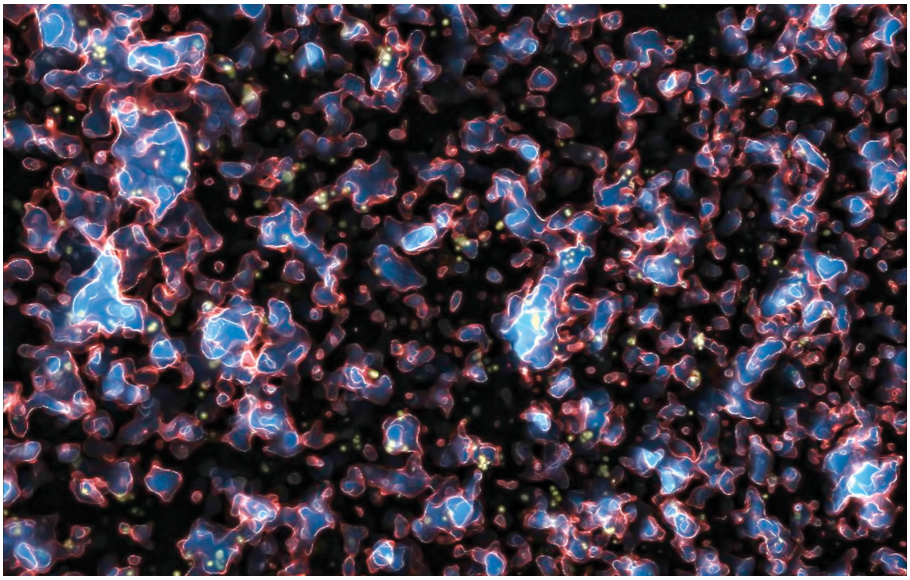
dark matter, nobody has identified exactly what it is. Signals from the EOR would help to indicate whether dark matter consists of relatively sluggish, or 'cold', particles — the model that is currently favoured — or 'warm' ones that are lighter and faster, says Anna Bonaldi, an astrophysicist at the Square Kilometre Array (SKA) Organisation near Manchester, UK. "The exact nature of dark matter is one of the things at stake," she says.

Although astronomers are desperate to learn more about the EOR, they are only now starting to close in on the ability to detect it. Leading the way are radio telescope arrays, which compare signals from multiple antennas to detect variations in the intensity of waves arriving from different directions in the sky.

One of the most advanced tools in the chase is the Low-Frequency Array (LOFAR), which is scattered across multiple European countries and centred near the Dutch town of Exloo. Currently the largest low-frequency radio observatory in the world, it has so far only been able to put limits on the size distribution of the bubbles, thereby excluding some extreme scenarios, such as those in which the intergalactic medium was particularly cold, says Leon Koopmans, an astronomer at the University of Groningen in the Netherlands who leads the EOR studies for LOFAR. Following a recent upgrade, a LOFAR competitor, the Murchison Widefield Array (MWA) in the desert of Western Australia, has further refined those limits in results due to be published soon.

In the short term, researchers say the best chance to measure the actual statistical properties of the EOR — as opposed to placing limits on them — probably rests with another effort called the Hydrogen Epoch of Reionization Array (HERA). The telescope, which consists of a set of 300 parabolic antennas, is being completed in the Northern Cape region of South Africa and is set to start taking data this month. Whereas the MWA and LOFAR are general purpose long-wavelength observatories, HERA's design was optimized for detecting primordial hydrogen. Its tight packing of 14-metre-wide dishes covers wavelengths from 50–250 MHz. In theory, that should make it sensitive to the cosmic-dawn trough, when galaxies first began to light up the cosmos, as well as to the EOR (see 'An Earth's-eye view of the early Universe').

As with every experiment of this kind, HERA will have to contend with interference from the Milky Way. The radio-frequency emissions from our Galaxy and others are thousands of times louder than the hydrogen line from the primordial Universe, cautions HERA's principal investigator, Aaron Parsons, a radioastronomer at the University of California, Berkeley. Fortunately, the Galaxy's emissions have a smooth, predictable spectrum, which can be subtracted to reveal cosmological features. To do so, however, radioastronomers must know exactly how their instrument responds to different wavelengths, also known as its systematics. Small changes in the surrounding environment, such



M. ALVAREZ, R. KAEHLER AND T. ABEL/ESO

A simulation of the epoch of reionization in the early Universe. Ionized material around new galaxies (bright blue) would no longer emit 21-centimetre radiation. Neutral hydrogen, still glowing at 21 cm, appears dark.

as an increase in soil moisture or pruning of a nearby bush, can make a difference — just as the quality of an FM radio signal can change depending on where you sit in a room.

If things go well, the HERA team might have its first EOR results in a couple of years, Parsons says. Nichole Barry, an astrophysicist at the University of Melbourne, Australia, and a member of the MWA collaboration, is enthusiastic about its chances: "HERA is going to have enough sensitivity that, if they can get the systematics under control, then boom! They can make a measurement in a short amount of time."

Similar to all existing arrays, HERA will aim to measure the statistics of the bubbles, rather than produce a 3D map. Astronomers' best hope for 3D maps of the EOR lie in the US$785-million SKA, which is expected to come online in the next decade. The most ambitious radio observatory ever, the SKA will be split between two continents, with the half in Australia being designed to pick up frequencies of 50–350 MHz, the band relevant to early-Universe hydrogen. (The other half, in South Africa, will be sensitive to higher frequencies.)

CRO-MAGNON COSMOLOGY

Although arrays are getting bigger and more expensive, another class of 21-cm projects has stayed humble. Many, such as EDGES, collect data with a single antenna and aim to measure some property of radio waves averaged over the entire available sky.

The antennas these projects use are "fairly Cro-Magnon", says CfA radioastronomer Lincoln Greenhill, referring to the primitive nature of the equipment. But researchers spend years painstakingly tweaking instruments to affect their systematics, or using computer models to work out exactly what the systematics are. This is a "masochistic obsession", says Greenhill, who leads the Large-Aperture Experiment to Detect the Dark Ages (LEDA) project in the United States. He often takes solo field trips to LEDA's antennas in Owens Valley, California, to do various tasks. These might include laying a new metal screen on the desert ground beneath the antennas, to act as a mirror for radio waves.

Such subtleties have meant that the community has been slow to accept the EDGES findings. The cosmic-dawn signal that EDGES saw was also unexpectedly large, suggesting that the hydrogen gas that was around 200 million years after the Big Bang was substantially colder than theory predicted, perhaps 4 kelvin instead of 7 kelvin. Since the release of the results in early 2018, theorists have written dozens of papers proposing mechanisms that could have cooled the gas, but many radioastronomers — including the EDGES team — warn that the experimental findings need to be replicated before the community can accept them.

LEDA is now attempting to do so, as are several other experiments in even more remote and inaccessible places. Ravi Subrahmanyan at the Raman Research Institute in Bengaluru, India, is working on a small, spherical antenna called SARAS 2. He and his team took it to a site on the Tibetan Plateau, and they are now experimenting with placing it on a raft in the middle of a lake. With fresh water, "you are assured you have a homogeneous medium below", Subrahmanyan says, which could make the antenna's response much simpler to understand, compared to that on soil.

Physicist Cynthia Chiang and her colleagues at the University of KwaZulu-Natal in Durban, South Africa, went even farther — halfway to Antarctica, to the remote Marion Island — to set up their cosmic-dawn experiment, called Probing Radio Intensity at High-Z from Marion. Chiang, who is now at McGill University in Montreal, Canada, is also travelling to a new site, Axel Heiberg Island in the Canadian Arctic. It has limited radio interference, and the team hopes to be able to detect frequencies as low as 30 MHz, which could allow them to detect the dark-ages trough.

At such low frequencies, the upper atmosphere becomes a serious impediment to observations. The best place on Earth to do them might be Dome C, a high-elevation site in Antarctica, Greenhill says. There, the auroras — a major source of interference — would be below the horizon. But others have their eyes set on space, or on the far side of the Moon. "It's the only radio-quiet location in the inner Solar System," says astrophysicist Jack Burns at the University of Colorado Boulder. He is leading proposals for a simple telescope to be placed in lunar orbit, as well as an array to be deployed by a robotic rover on the Moon's surface.

Other, more conventional techniques have made forays into the first billion years of the Universe's history, detecting a few galaxies and quasars — black-hole-driven beacons that are among the Universe's most luminous phenomena. Future instruments, in particular the James Webb Space Telescope that NASA is due to launch in 2021, will bring more of these findings. But for the foreseeable future, conventional telescopes will spot only some of the very brightest objects, and therefore will be unable to do any kind of exhaustive survey of the sky.

The ultimate dream for many cosmologists is a detailed 3D map of the hydrogen not only during the EOR, but all the way back to the dark ages. That covers a vast amount of space: thanks to cosmic expansion, the first billion years of the Universe's history account for 80% of the current volume of the observable Universe. So far, the best 3D surveys of galaxies — which tend to cover closer, and thus brighter, objects — have made detailed maps of less than 1% of that volume, says Max Tegmark, a cosmologist at the Massachusetts Institute of Technology in Cambridge. Loeb, Tegmark and others have calculated that the variations in hydrogen density before the EOR contain much more information than the CMB does[3,4], which so far has been the gold standard for measuring the main features of the Universe. These include its age, the amount of dark matter it contains and its geometry.

Mapping this early hydrogen will be a huge technical challenge. Jordi Miralda-Escudé, a cosmologist at the University of Barcelona in Spain, says that with current technology, it is so challenging as to be a "pipe dream".

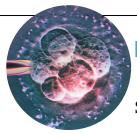
But the pay-off of producing such maps would be immense, says Loeb. "The 21-cm signal offers today the biggest data set on the Universe that will ever be accessible to us." ■

Davide Castelvecchi *is a senior reporter for* Nature *based in London.*

1. Bowman, J. D., Rogers, A. E. E., Monsalve, R. A., Mozdzen, T. J. & Mahesh, N. *Nature* **555**, 67–70 (2018).
2. Zeldovich, Y. B., Kurt, V. G. & Syunyaev, R. A. [in Russian] *Z. Eksp. Teor. Fiz.* **55**, 278–286 (1968).
3. Loeb, A. *J. Cosmol. Astropart. Phys.* **2012**, 028 (2012).
4. Mao, Y. Tegmark, M., McQuinn, M. Zaldarriaga, M. & Zahn, O. *Phys. Rev. D* **78**, 023529 (2008).

COMMENT

POLICY Why bioethics is continually outpaced by technology **p.308**



HISTORY The mosquito's outsized role in conflicts since antiquity **p.310**

SUSTAINABILITY Sand plunder risks health and squanders buffer to sea-level rise **p.312**

SOCIAL JUSTICE Astronomer says stop and listen to Native Hawaiians at Mauna Kea **p.312**

VASSIL DONEV/EPA/SHUTTERSTOCK

A child holds a sign protesting against genetically modified crops during a demonstration in Bulgaria.

Key concepts for making informed choices

Teach people to think critically about claims and comparisons using these concepts, urge **Andrew D. Oxman** and an alliance of 24 researchers — they will make better decisions.

Everyone makes claims about what works. Politicians claim that stop-and-search policing will reduce violent crime; friends might assert that vaccines cause autism; advertisers declare that natural food is healthy. A group of scientists describes giving all school-children deworming pills in some areas as one of the most potent anti-poverty interventions of our time. Another group counters that it does not improve children's health or performance at school.

Unfortunately, people often fail to think critically about the trustworthiness of claims, including policymakers who weigh up those made by scientists. Schools do not do enough to prepare young people to think critically[1]. So many people struggle to assess evidence. As a consequence, they might make poor choices.

To address this deficit, we present here a set of principles for assessing the trustworthiness of claims about what works, and for making informed choices (see ▶

KEY CONCEPTS FOR INFORMED CHOICES

This framework assists people helping others to think critically and make informed decisions.

CLAIMS: Claims about effects should be supported by evidence from fair comparisons. Other claims are not necessarily wrong, but there is an insufficient basis for believing them.

Claims should not assume that interventions are safe, effective or certain.
- Interventions can cause harm as well as benefits.
- Large, dramatic effects are rare.
- We can rarely, if ever, be certain about the effects of interventions.

Seemingly logical assumptions are not a sufficient basis for claims.
- Beliefs alone about how interventions work are not reliable predictors of the presence or size of effects.
- An outcome may be associated with an intervention but not caused by it.
- More data are not necessarily better data.
- The results of one study considered in isolation can be misleading.
- Widely used interventions or those that have been used for decades are not necessarily beneficial or safe.
- Interventions that are new or technologically impressive might not be better than available alternatives.
- Increasing the amount of an intervention does not necessarily increase its benefits and might cause harm.

Trust in a source alone is not a sufficient basis for believing a claim.
- Competing interests can result in misleading claims.
- Personal experiences or anecdotes alone are an unreliable basis for most claims.
- Opinions of experts, authorities, celebrities or other respected individuals are not solely a reliable basis for claims.
- Peer review and publication by a journal do not guarantee that comparisons have been fair.

COMPARISONS: Studies should make fair comparisons, designed to minimize the risk of systematic errors (biases) and random errors (the play of chance).

Comparisons of interventions should be fair.
- Comparison groups and conditions should be as similar as possible.
- Indirect comparisons of interventions across different studies can be misleading.
- The people, groups or conditions being compared should be treated similarly, apart from the interventions being studied.
- Outcomes should be assessed in the same way in the groups or conditions being compared.
- Outcomes should be assessed using methods that have been shown to be reliable.
- It is important to assess outcomes in all (or nearly all) the people or subjects in a study.
- When random allocation is used, people's or subjects' outcomes should be counted in the group to which they were allocated.

Syntheses of studies should be reliable.
- Reviews of studies comparing interventions should use systematic methods.
- Failure to consider unpublished results of fair comparisons can bias estimates of effects.
- Comparisons of interventions might be sensitive to underlying assumptions.

Descriptions should reflect the size of effects and the risk of being misled by chance.
- Verbal descriptions of the size of effects alone can be misleading.
- Small studies might be misleading.
- Confidence intervals should be reported for estimates of effects.
- Deeming results to be 'statistically significant' or 'non-significant' can be misleading.
- Lack of evidence for a difference is not the same as evidence of no difference.

CHOICES: What to do depends on judgements about the problem, the relevance (applicability or transferability) of evidence available and the balance of expected benefits, harm and costs.

Problems, goals and options should be defined.
- The problem should be diagnosed or described correctly.
- The goals and options should be acceptable and feasible.

Available evidence should be relevant.
- Attention should focus on important, not surrogate, outcomes of interventions.
- There should not be important differences between the people in studies and those to whom the study results will be applied.
- The interventions compared should be similar to those of interest.
- The circumstances in which the interventions were compared should be similar to those of interest.

Expected pros should outweigh cons.
- Weigh the benefits and savings against the harm and costs of acting or not.
- Consider how these are valued, their certainty and how they are distributed.
- Important uncertainties about the effects of interventions should be reduced by further fair comparisons.

'Key Concepts for Informed Choices'). We hope that scientists and professionals in all fields will evaluate, use and comment on it. The resources were adapted, drawing on the expertise of two dozen researchers, from a framework developed for health care[2] (see 'Randomized trial').

Ideally, these concepts should be embedded in education for citizens of all ages. This should be done using learning resources and teaching strategies that have been evaluated and shown to be effective.

TRUSTWORTHY EVIDENCE

People are flooded with information. Simply giving them more is unlikely to be helpful, unless its value is understood. A 2016 survey in the United Kingdom showed that only about one-third of the public trusts evidence from medical research; about two-thirds trust the experiences of friends and family[3].

Not all evidence is created equal. Yet people often don't appreciate which claims are more trustworthy than others; what sort of comparisons are needed to evaluate different proposals fairly; or what other information needs to be considered to inform good choices.

For example, many people don't grasp that two things can be associated without one

necessarily causing the other. The media sometimes perpetuates this problem by using language suggesting that cause and effect has been established when it has not[4] — for instance, statements such as 'coffee can kill you' or 'drinking one glass of beer a day can make you live longer'. Worse, exaggerated causal claims often pepper press releases from universities and journals[5].

Studies that make fair comparisons are crucial, yet people often don't know how to appraise the validity of research. Systematic reviews that synthesize well-designed studies that are relevant to clearly defined questions are more trustworthy than haphazard observations. This is because they are less susceptible to biases (systematic distortions) and the play of chance (random errors). Yet results from single studies are often reported in isolation, as facts. Hence the familiar flip-flopping headlines such as 'chocolate is good for you', followed the next week by 'chocolate is bad for you'.

To make good choices, other types of information are needed too — for example, about costs and feasibility. Judgements must also be made about the relevance of information from research (how applicable or transferable it is), and about the balance between the likely desirable and undesirable effects of a drug, therapy or regulation.

When it comes to carbon taxes, for example, policymakers need to consider evidence about the environmental and economic effects of such taxes, judge how comparable their context is with that of the studies and weigh how onerous the administrative difficulties are. They also need to model how tax burdens will be distributed across socio-economic groups and think about whether the taxes will be accepted in their jurisdictions.

CRITICAL THINKING

Individuals and organizations across many fields are working to enable people to make informed decisions. These efforts include synthesizing the best available evidence in systematic reviews; making that information more accessible, such as through plain-language summaries or open access; and teaching people how to use such resources. Examples of such review organizations are Cochrane (previously called the Cochrane Collaboration), which focuses on health care; the Campbell Collaboration, which looks at the effects of social policies; the Collaboration for Environmental Evidence; and the International Society for Evidence-Based Health Care. Others include the Center for Evidence-Based Management, the Africa Centre for Evidence, the International Initiative for Impact Evaluation (known as 3ie) and Britain's What Works Centres.

Unfortunately, academics tend to work in silos and can miss opportunities to learn from others. The expertise of the authors of this article spans 14 fields: agriculture, economics, education, environmental management, international development, health care, informal learning, management, nutrition, planetary health, policing, speech and language therapy, social welfare, and veterinary medicine.

We have identified many concepts that apply across these fields (see 'Key Concepts for Informed Choices' and 'Key concepts in action'). Some further concepts are more relevant in some fields than in others. For example, it is often important to consider potential placebo effects when assessing claims about medical treatments and nutrition; these are rarely relevant to interventions in the environment.

Our collaboration has already prompted many of us to develop frameworks for specific fields and to suggest improvements to the original Informed Health Choices framework[2]. There is power in identifying an issue that resonates across different domains; it provides momentum to align efforts.

The Key Concepts for Informed Choices is not a checklist. It is a starting point. Although we have organized the ideas into three groups (claims, comparisons and choices), they can be used to develop learning resources that include any combination of these, presented in any order. We hope that the concepts will prove useful to people who help others to think critically about what evidence to trust and what to do, including those who teach critical thinking and those responsible for communicating research findings.

NEXT STEPS

Evidence-informed practice is now taught to professionals in many different fields, and these efforts must grow. It is also crucial that

MIKKEL OSTERGAARD/PANOS

Pupils at a school in Uganda.

RANDOMIZED TRIAL

Children taught key concepts pass test

The Informed Health Choices (IHC) Project was initially developed between 2012 and 2017 by a collaboration including some of the co-authors of this article (A.D.O., A.D., I.C. and M.O.). The project includes its own set of key concepts[2], learning resources and a database of multiple-choice questions to assess how well users can apply the concepts.

In 2016, a randomized trial involving 120 schools and more than 10,000 schoolchildren in Uganda showed that these resources improved the ability of 10–12-year-old children to apply 12 of the key concepts[7]. These concepts included, for example, recognizing that personal experiences alone are an insufficient basis for claims about effects, and that small studies can be misleading.

In this trial, 69% of schoolchildren who were taught the key concepts passed a multiple-choice test of their ability to think critically about health claims. By comparison, just 27% of children who were not told about the concepts passed the same test. A.D.O. *et al.*

KEY CONCEPTS IN ACTION

Examples of evaluated evidence

CLAIMS

Beliefs alone about how interventions work are not reliable predictors of the presence or size of effects.

Most people feel that it is hard to influence parents' engagement with their children's education. The assumption is therefore that more intensive (and more costly) interventions would be more likely to be effective. However, studies of intensive interventions have often failed to show effects on pupils' attainment, as measured using standard tests (see go.nature.com/2gfy8io).

Meanwhile, a recent evaluation of the effects of simply text-messaging parents weekly with updates about their child's schooling had positive effects on children's attendance, homework submission and mathematics attainment (see go.nature.com/2t7ormy). These effects were small, but the cost was very low. This illustrates that — contrary to our hunches — inexpensive interventions can be helpful, and expensive ones can fail.

A maternity ward in Dar es Salaam, Tanzania.

GARY CARLTON/ALAMY

COMPARISONS

Conditions should be as similar as possible.

'Scared Straight' programmes take young offenders on prison visits on the assumption that this experience and listening to inmates' descriptions of life inside will deter juvenile delinquency. Some studies have found that such prison visits were followed by large reductions in delinquent behaviour. But a lot can change in a group of youngsters over time, including their becoming older and more mature. How can anyone know that the prison visits caused the reduction?

Fairer experiments were done in which youths were randomly assigned to visit prison or not, creating groups that were more comparable. Comparisons between these groups showed more delinquency in the youngsters who had been exposed to prisons than among those who had not[8,9].

CHOICES

When there are important uncertainties about the effects of interventions, those should be reduced by fair comparisons.

In the health sector, financing schemes in which funds are released only if a specific action is taken or performance target is met have become popular. Billions of dollars have been invested in promoting these schemes in low- and middle-income countries, with the aim of achieving international development goals[10]. For example, health providers have been offered cash rewards for increasing the percentage of births in clinics (rather than at home), with the intention of improving maternal and newborn health and survival.

But performance-based financing schemes can have unintended adverse effects, such as encouraging health-care workers to falsify records or to neglect other activities. In Tanzania, some health facilities threatened new mothers with fines or denial of vaccinations for their children[10]. For interventions in which there is much uncertainty about the pros and cons, further fair comparisons should be done before or while rolling out such schemes. **A.D.O. *et al.***

schoolchildren learn these key concepts, rather than delaying acquisition of these skills until adulthood. Young people who have been explicitly taught critical thinking make better judgements than those who have not[6]. Educating people about such concepts at a young age sets an important foundation for future learning.

An important part of the work of encouraging critical thinking is learning and sharing strategies that promote healthy scepticism, but which avoid unintended adverse consequences. These include inducing nihilism (extreme scepticism); allowing for disingenuous claims that uncertainty is a defensible argument against action (on climate change, for example); or encouraging false beliefs — such as that all research is untrustworthy because of competing interests among those who promote particular interventions.

Competing interests take various forms in different fields, but the challenges and remedies are similar: recognition of potential conflicts, transparency and independent evaluations. Achieving these depends on improved public understanding of the need for independent evaluation, and public demand for investment in it, as well as unbiased communication of findings.

Further development and specialization of the Key Concepts for Informed Choices is needed, and we welcome suggestions. For example, more consideration needs to be given to how these concepts can be applied to actions to address system-wide changes, taking into account complex, dynamic interactions and feedback loops, such as in climate-change mitigation or adaptation strategies.

We have therefore created a website (www.thatsaclaim.org) on which our key concepts can be adapted to different fields and target users, translated into other languages and linked to learning resources. ■

Andrew D. Oxman *is research director at the Centre for Informed Health Choices, Norwegian Institute of Public Health, Oslo, Norway.* **Jeffrey K. Aronson, Eric Barends, Robert Boruch, Marnie Brennan, Iain Chalmers, Joe Chislett, Peter Cunliffe-Jones, Astrid Dahlgren, Marie Gaarder, Andy Haines, Carl Heneghan, Robert Matthews, Brandy Maynard, Matt Oxman, Andrew Pullin, Nicola Randall, Hazel Roddam, Anel Schoonees, Jonathan Sharples, Ruth Stewart, Janet Stott, Raymond Tallis, Nerys Thomas, Luke Vale.**
e-mail: oxman@online.no

1. Bouygues, H. L. *The State of Critical Thinking: A New Look at Reasoning at Home, School, and Work* (Reboot Foundation, 2018).
2. Oxman, A. D., Chalmers, I., Austvoll-Dahlgren, A. & Informed Health Choices group. *F1000Research* **7**, 1784 (2018).
3. Academy of Medical Sciences. *Enhancing the Use of Scientific Evidence to Judge the Potential Benefits and Harms of Medicines* (Academy of Medical Sciences, 2017).
4. Haber, N. *et al. PLoS ONE* **13**, e0196346 (2018).
5. Sumner, P. *et al. PLoS ONE* **11**, e0168217 (2016).
6. Abrami, P. C. *et al. Rev. Educ. Res.* **85**, 275–314 (2015).
7. Nsangi, A. *et al. Lancet* **390**, 374–388 (2017).
8. Petrosino, A., Turpin-Petrosino, C. & Finckenauer, J. O. *Crime Delinq.* **46**, 354–379 (2000).
9. Petrosino, A., Turpin-Petrosino, C., Hollis-Peel, M. E. & Lavenberg J. G. *Cochrane Database Syst. Rev.* CD002796 (2013).
10. Renmans, D., Holvoet, N., Orach, C. G. & Criel, B. *Health Pol. Plan.* **31**, 1297–1309 (2016).

A full list of author affiliations accompanies this Comment online (see go.nature.com/2tdzvn7).

BILL CLARK/CQ ROLL CALL

Calls for universal US health coverage under Medicare face challenges, such as the fact that end-stage kidney disease consumes 7% of its budget.

HEALTH CARE

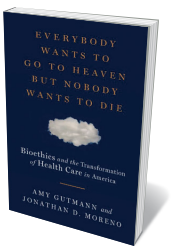
The battle in bioethics

Eric J. Topol weighs up a book on a field sprinting to keep up with biotechnology.

The term 'bioethics' was coined in 1926, yet the field itself did not emerge until the 1970s. Although my 1975 university thesis (Prospects for Genetic Therapy in Man) reviewed ethical concerns, it took a further four decades before gene therapy was successful in people. More recently, some developments in biomedical technology have accelerated beyond moral or principled boundaries. Among the most shocking was last November's revelation that the premature and reckless application of human-embryo genome editing had given rise to twin babies in China. That led to calls for a global moratorium (see *Nature* **566**, 440–442; 2019).

Amy Gutmann and Jonathan Moreno have long been at the heart of bioethics debates, and served together for seven years on Barack Obama's Presidential Commission for the Study of Bioethical Issues. Their book *Everybody Wants to Go to Heaven but Nobody Wants to Die* (its title is borrowed from a country-music song) reviews the field's evolution and status.

To begin, Gutmann and Moreno each recount a personal flashback to an older era of ethically problematic medical care. Gutmann's grandmother and Moreno's mother underwent medical amputations; neither had been given crucial information by her doctors, so both were uninformed at the time of crucial therapeutic decisions. The authors then tour ethical dilemmas throughout the human life cycle, ranging from reproductive rights to the right to die.



Everybody Wants to Go to Heaven but Nobody Wants to Die: Bioethics and the Transformation of Health Care in America
AMY GUTMANN & JONATHAN D. MORENO
Liveright (2019)

Many of the stops along the way delve into familiar territory — required reading for clinical researchers, and the basis of annual online-testing requirements for conducting clinical research. For instance, the authors cover the infamous, decades-long Tuskegee syphilis study, in which the US Public Health Service withheld penicillin from hundreds of African Americans with the illness. And they discuss the case of Jesse Gelsinger, who died in 1999 from misguided gene therapy intended to treat the rare metabolic disorder ornithine transcarbamylase deficiency.

The authors are not shy about expressing their liberal views, many of which I share. For instance, they declare that health care is a human right, and they believe that people should have the freedom to access safe and legal abortions.

Against a background of calls for "Medicare for All" by several Democratic Party presidential hopefuls, Gutmann and Moreno discuss this government-run,

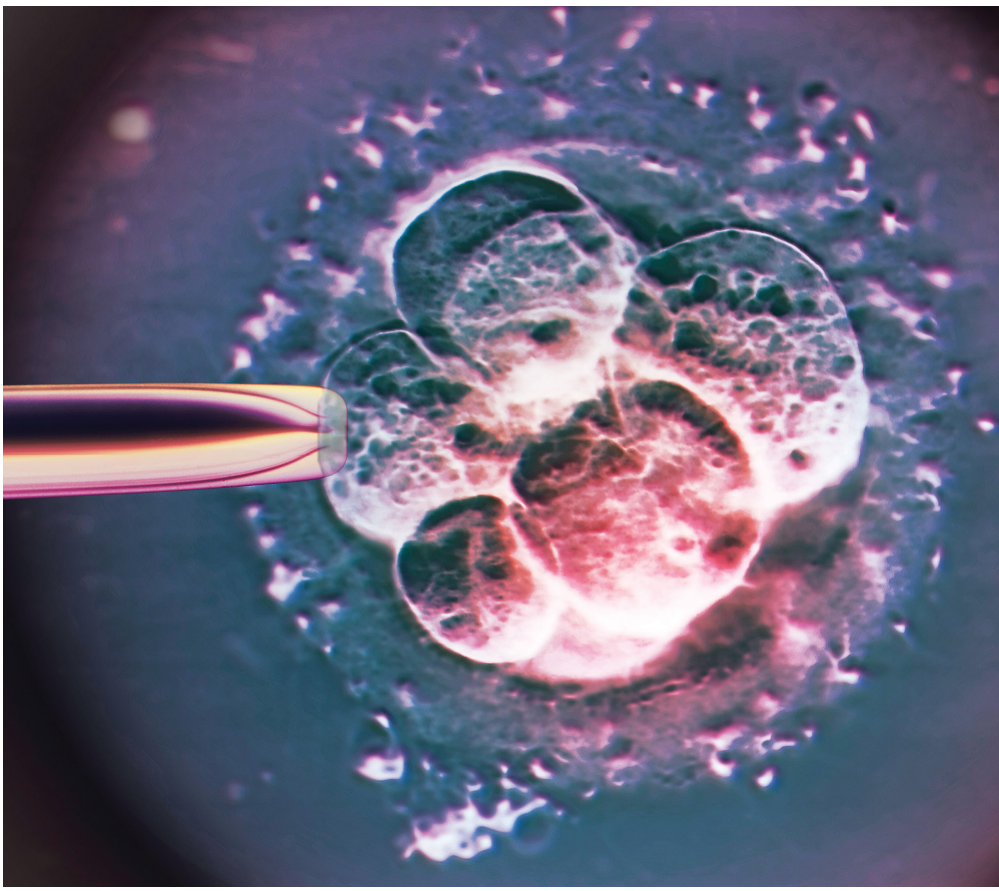
taxpayer-funded health insurance scheme. Covering US citizens aged 65 and older and signed into law in 1965, Medicare was extended in 1972 to include all people with end-stage kidney disease, irrespective of age or demographic. The full costs of dialysis are now footed for more than 500,000 US citizens at a cost of more than US$30 billion a year. And care for end-stage kidney disease consumes approximately 7% of the Medicare budget.

This federal carve-out has fuelled for-profit dialysis centres nationwide. Ultimately, it has caused a lack of financial support for an untold number of people with other conditions, including some with haemophilia or with one of many rare diseases for which treatments are costly and often involve injectable speciality drugs. This demonstrates the problem of providing health care for everyone affected by just one condition, as well as the economic implications of coverage for all in a country that has the highest medical expenditure per person in the world.

Despite their vast experience and wisdom, the authors make important errors. One pertains to mitochondrial-replacement therapy. The powerhouses of our cells, mitochondria contain only 0.1% of our DNA, but mutations in that genetic material (known as mtDNA) can be the root cause of rare diseases transmitted from mother to child. To counter this potential when a prospective mother has such mutations, another woman without the mutation can provide donor mtDNA amounting to 0.0005% of the embryo's genome. Gutmann and Moreno write that, in 2016, the United States gave the green light for male embryos to be given the treatment. In fact, the procedure is still banned by the US Food and Drug Administration, although Britain legalized it in 2015. The authors also erroneously indicate that angiograms — X-rays of blood vessels — can support diagnoses of brain death in people in persistent vegetative states.

A major theme throughout is that patients have more agency and authority today than they once did, and can even co-produce their care, sharing key decisions with their doctors. But the authors' proclamation that there has been "a collapse of medical paternalism" is off-base. Unfortunately, paternalism is still pervasive. As I noted in my 2014 book, *The Patient Will See You Now*, some 66% of US doctors will not give patients their office notes, and almost all order routine medical scans without telling the recipient how much exposure to ionizing radiation the tests entail.

There are also key omissions. I was surprised to see no mention of non-invasive prenatal tests, which have accurately identified the potential for fetal chromosomal abnormalities for more than one million prospective parents in the United States. They do not discuss ongoing clinical trials using induced pluripotent stem cells to treat medical conditions such as macular degeneration, Parkinson's disease or spinal injury. They barely mention the 'brain in a dish' approach to neurological research involving human cortical organoids, which is attracting considerable attention from bioethicists. And devoting just a handful of sentences to CRISPR genome editing of human embryos and subsequent births seems remiss.



ZEPHYR/SPL

Replacing faulty mitochondrial DNA in embryos is allowed under UK, but not US, law.

> ***"Patients have more agency and authority today than they once did, and can even co-produce their care."***

Nor do they mention one of the most controversial bioethics incidents in recent years. In 2015, the cognitive psychologist Steven Pinker wrote in the newspaper *The Boston Globe*: "Biomedical research will always be closer to Sisyphus than a runaway train — and the last thing we need is a lobby of so-called ethicists helping to push the rock down the hill." Inevitably, bioethicists pushed back at this declaration that they are a kind of guild, a bureaucratic industry entangled in a conflict of interest. It's a shame that Gutmann and Moreno don't tackle this frontal assault. The moral compass that bioethicists provide is necessary: all too often, technology is out in front of the deep thinking we need about how it can be best applied.

Indeed, bioethics is often pivotal in educating clinicians about patient care at academic medical centres. That brings me to the concept of casuistry: thinking about ethical problems by assessing a spectrum of cases to which they apply. The book stresses that careful analysis of a case can promote insight.

I experienced this at first hand on my rounds as attending physician in an intensive-care unit. I and my team of medical students and trainees cared for many people facing death. We had to consider 'do not resuscitate' orders, and discovered how best to discuss the delicate situation with patients and their families. No one was more thoughtful while weighing in than the bioethicists. When they were absent, there was a sense of loss: we missed their clarity. Whether in the context of an individual patient, a medical-research initiative or the application of new advances, the field of bioethics is essential. We will continue to rely on these professionals for guidance. ■

Eric J. Topol *is professor of molecular medicine at Scripps Research in La Jolla, California.*
e-mail: etopol@scripps.edu

SARAH L. VOISIN/WASHINGTON POST/GETTY

A multiple-exposure image of a single *Anopheles punctipennis* mosquito, on display at the Smithsonian Natural History Museum in Washington DC.

MEDICAL HISTORY

Murderous trail of the mosquito

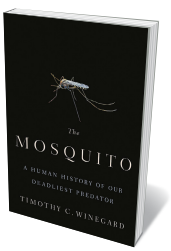
Karen Masterson appraises the disease vector's role in scientific and military history.

The deadliest beast on Earth is the featherweight mosquito. Among the diseases it passes on — such as filariasis, yellow fever, dengue, Zika and West Nile fever — malaria accounted for 435,000 deaths in 2017. Inevitably, the insect has inspired many books. So what does the baldly titled tome *The Mosquito* add to the canon? The answer: a lot.

Military historian Timothy Winegard's book takes readers on a riveting adventure, documenting the mosquito's outsized role in conflict since antiquity. He shows how, from vast empires to contemporary war zones, the advantage fell to any defending army able to stall attackers in mosquito-filled swamps, where fevers — mostly malaria, yellow fever and dengue — sapped their strength. Through this lens, he explains in superb detail how great powers rose and fell.

We learn how, in fifth-century BC Greece, Persian troops crumbled when a coalition of Athenians and Spartans forced them into marshlands before the Battle of Plataea. A mix of malaria (transmitted by the *Anopheles* mosquito) and dysentery felled 40% of the Persian ranks. Thus "General Anopheles", writes Winegard, freed Greeks from Persian rule, and enabled the blossoming of Greek philosophy, science and art — the 'Golden Age' that in part paved the way for Western civilization.

Winegard relates, too, the mosquito's role in the fall of Rome. For hundreds of years up to the fifth century AD, the malarial Pontine Marshes around Rome staved off attacks by Carthaginians, Germanic tribes and Huns, yet weakened Roman citizens. Over the following 300 years, malaria also helped to ground the Holy Roman Empire. Christian hospitals took in the masses of infected people, proselytizing a doctrine of care that won over pagans and ultimately paved the way for Charlemagne's claim on Europe. The thread of influence runs all the way to the Vietnam War, when mosquito-borne diseases made US occupation of the North Vietnamese-held jungles untenable.



The Mosquito: A Human History of Our Deadliest Predator
TIMOTHY C. WINEGARD
Dutton (2019)

Although Winegard's approach is at times too broad and unscientific, it is fascinating. And he covers some research well, such as why mosquitoes bite humans selectively. (The science is still uncertain, but evidence suggests that 20% of people receive 80% of bites: T. A. Perkins *et al. PLoS Comput. Biol.* **9**, e1003327; 2013.) Ultimately, however, Winegard is strongest on the world-changing aspects of malaria — not only the rise and fall of empires, but also areas such as the nexus of genetics, society and politics.

He discusses, for instance, a link between the Atlantic slave trade and genetic resistance to the malaria pathogen *Plasmodium vivax*. The Duffy antigen on red blood cells is the receptor for *P. vivax*, thus helping to launch infection. Anyone lacking this antigen is resistant to malaria, and that includes 95% of people of West African descent. Most of the people wrenched into slavery in the Americas came from this region; and, in a horrible irony, their resistance to the disease — observed by plantation owners — created demand for their work and became a driver of the slave trade.

Winegard overflows with enthusiasm for his subject. At times, however, his hammering at perspectives that he wants us to take away comes at the expense of nuance and specificity. A case in point is his narrative on a horrific chapter of the Second World War, when human experiments to find an antimalarial were carried out by both Nazi Germany and the United States.

Winegard cites my work (*The Malaria Project*, 2014), which details how the German malariologist Claus Schilling deliberately infected some 1,000 prisoners in the Dachau concentration camp. Winegard claims that 400 of them perished as a result. My research showed that 38 died in Schilling's hospital wing, from the effects of two particularly toxic drugs. Typhus and dysentery might well have killed others, once they were released back to the barracks.

Meanwhile, Winegard downplays the fact that more than 100 US doctors were simultaneously doing the same thing, on a vastly greater scale. They experimented on 10,000 enlisted military personnel and inmates at six state hospitals and three prisons — including the notorious Stateville Penitentiary outside Chicago, Illinois. The death toll

is estimated to have been between 10% and 30%. By not including this shocking episode of US medical history, Winegard misses a valuable application of his own framework: that one insect had driven so many doctors into inhumane, medically desperate work.

Winegard also lacks nuance in asserting that new technologies will soon extinguish the mosquito. Some of the science is indeed promising. Research published this month in *Nature*, for instance, shows how irradiation and bacterial infection were used to nearly eradicate tiger mosquitoes (*Aedes albopictus*) from two river islands in China (X. Zheng *et al. Nature* **572**, 56–61; 2019). Mosquito specialist Peter Armbruster has questioned the scalability and sustainability of this work, however (P. A. Armbruster *Nature* **572**, 39–40; 2019) . Even trickier is the use of CRISPR and gene drives, in which laboratory-raised mosquitoes pass on an altered gene to generations of wild mosquitoes — work that is still theoretical.

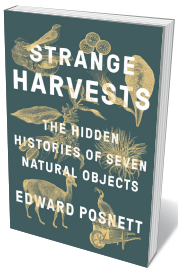
Finally, there is the issue of Winegard's view of the insect. He sees it solely in human terms, setting it up as a foe to humanity, with no other role in nature. To claim that some 3,000 species of mosquito exist for no reason other than to act as our "apex predator" is bold, but it rests on unsound scientific footing.

His larger points in *The Mosquito* remain valuable, built on the solid work of scholars and scientists. Ever since *Homo sapiens* moved away from hunting and gathering, we have paid dearly for tangling with nature. As we tear down forests, cultivate fields and transform our environment, we create perfect habitats for mosquito propagation. And — more to Winegard's point — as we shred the Earth with weaponry and park large armies in marshlands, we create ideal conditions for mosquitoes to spread disease. Humans, he rightly notes, help mosquito species to diversify, adapt and thrive as we reshape the planet.

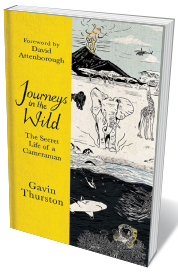
When mosquitoes turn to us for blood, they transfer all the microbes they've evolved to carry. We have had no choice but to fight back. So, in this sense, we are at war with mosquitoes, from the multi-billion-dollar global health campaigns against mosquito-borne diseases — funded largely by wealthy countries through international agreements, and donors such as the Bill & Melinda Gates Foundation — to the pesticides that many spray in their backyards. Mosquitoes control our behaviour because we have yet to control them. Winegard's earnest voice on this brings the seriousness of research and action on the mosquito up to the needed decibel. ■

Karen Masterson *is a science journalism professor at Stony Brook University and author of* The Malaria Project, *a narrative history of the US government's campaign to stop malaria during the Second World War. e-mail: karen.masterson@stonybrook.edu*

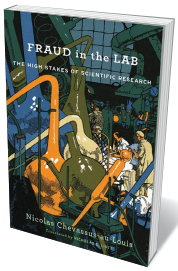
Books in brief



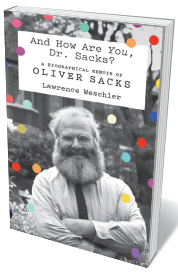
Strange Harvests
Edward Posnett VIKING *(2019)*
The global journeys of commodities such as salt have largely been told. In this subtle, reflective study, nature writer Edward Posnett follows the wake of seven very different products. Harvested from living plants and animals, they range from tagua (vegetable ivory, the nut of South American palm *Phytelephas*) to byssus (the 'sea silk' exuded by marine molluscs as anchorage). Woven through are moving stories of the remote microeconomies engaged in these trades, such as Iceland's eiderdown gatherers who, year on year, give safe haven to thousands of wild eider ducks in nesting season.



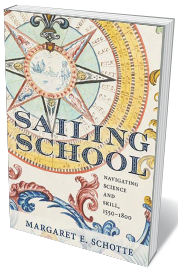
Journeys in the Wild
Gavin Thurston SEVEN DIALS *(2019)*
Neither plane crashes, political coups nor a mighty slap from a silverback gorilla have put wildlife cameraman Gavin Thurston off his stride. A force behind documentaries such as David Attenborough's BBC series *Blue Planet II*, Thurston has chased fauna worldwide for 40 years. His no-holds-barred memoir plunges you into the serendipities and perils of working in the remote wilderness, as he stands stock-still to 'hide' from short-sighted African elephants in Kenya, films demoiselle cranes flying 6 kilometres up above Nepal, or marvels at the hiss of Mauritania's dryland crocodiles.



Fraud in the Lab
Nicolas Chevassus-au-Louis, tr. Nicholas Elliott HARVARD UNIV. PRESS *(2019)*
This bracing critical analysis, now in its first English edition, skewers the 'publish or perish' lab culture driving scientific fraud. Science writer Nicolas Chevassus-au-Louis explores the terrain through cases such as medical researcher William Summerlin, who inked transplanted mouse skin to falsify results in the 1970s. And he shows the serious, real-life impacts of "data beautification", manipulated images and plagiarism. His solution for science? Think communally, end the tyranny of impact factors — and slow down.



And How Are You, Dr Sacks?
Lawrence Weschler FARRAR, STRAUS & GIROUX *(2019)*
In the 1980s, Oliver Sacks regularly met with journalist Lawrence Weschler for what became a four-year interview, casting back over the neurologist's tumultuous early career. That trove forms the bulk of Weschler's engrossing biographical memoir. This is Sacks at full blast: on endless ward rounds, observing his post-encephalitic patients (portrayed in his 1973 book *Awakenings*), exulting over horseshoe crabs and chunks of Iceland spar. Weschler ends by speculating that Sacks altered neurological practice itself through his attentive compassion for the patients who feature in his stories.



Sailing School
Margaret E. Schotte JOHNS HOPKINS UNIV. PRESS *(2019)*
From the Renaissance to the Enlightenment, a singular publishing boom played out in Europe's maritime nations. As voyages stretched into open ocean, mathematical expertise in celestial navigation became essential. Hands-on instruction with instruments remained key, but as historian Margaret Schotte reveals in this deft, scholarly chronicle, the nautical manual soon came into its own. Between 1509 and 1800, some 600 were published across 6 countries to impart the necessary theory, helping sailors to become scientists in the classroom as well as on ship's deck. **Barbara Kiser**

Correspondence

Danger to science of no-deal Brexit

As UK-based European stakeholders, we are deeply concerned about the threat that Brexit — particularly a 'no deal' scenario — poses to international research (*Nature* **572**, 13–14; 2019).

Uncertainties arising from the 2016 Brexit referendum have already undermined the attraction for foreigners of doing research in Britain. In our view, the various scenarios are all likely to damage research initiatives.

'Shadow membership' and 'third country' scenarios, for example, represent different degrees of cooperation with the European Union. These could introduce new challenges, and perhaps opportunities, with regard to partnerships, taxes and regulations. But they would still curtail the freedom enjoyed by European academics. The UK government would need to increase its research budget to offset the loss of the EU funding.

Scientific excellence is underpinned by researcher mobility, adequate resources and regulations that foster long-term stability and planning. A no-deal scenario would result in fewer European collaborations, diminished resources and constrained legal frameworks. It would therefore present a grave danger to science.

Mariana Pinto da Costa* *Queen Mary University of London, UK.*
mariana.pintodacosta@qmul.ac.uk
**On behalf of 4 correspondents; see go.nature.com/33rpv9j.*

Astronomy's ethical duty to Hawaiian site

As a partner of the Thirty Meter Telescope (TMT) Consortium and a member of Qalipu Mi'kmaq First Nations, I am one of very few Indigenous faculty members in Canadian astronomy. In my view, Canada's astronomy community has an ethical duty to listen to the Native Hawaiian protectors of the sacred Mauna Kea site, where the consortium now has a permit for construction. Our response will affect the future of astronomy and reconciliation with Indigenous peoples.

According to the United Nations Declaration on the Rights of Indigenous Peoples (UNDRIP) and the Calls of Action of the Truth and Reconciliation Commission of Canada, Canada and the TMT consortium have a duty to respect the wishes of the protectors, along with Indigenous peoples' rights, wherever we pursue astronomical discovery. The astronomy community should therefore halt construction, listen to the protectors and support those protesters who have been arrested.

If the consortium is not willing to step back, then Canada must remove itself from the project as part of its commitment to UNDRIP. Otherwise, we continue to support a culture that does not respect the right of self-determination and is not inclusive of Indigenous peoples.

Hilding Neilson *University of Toronto, Canada.*
hilding.neilson@utoronto.ca

Sand: an overlooked occupational hazard

Mette Bendixen and colleagues point out the environmental, social and economic harms that sand extraction might cause (*Nature* **571**, 29–31; 2019). It can also affect human health, a particularly important point for workers. A global agenda for sustainable sand extraction should incorporate workers' health policies to prevent silicosis and other serious lung diseases.

The surface properties that make sand from deserts or beaches unsuitable for the building industry also make it less hazardous when inhaled by humans. However, long-term inhalation of small crystalline particles of silica (sand's primary component) can lead not just to silicosis, a progressive and incurable fibrotic lung disease, but to lung cancer, chronic obstructive pulmonary disease, autoimmune disease and tuberculosis (P. Cullinan *et al. Lancet Respir. Med.* **5**, 445; 2017). Hazardous jobs that involve exposure to freshly fractured silica include crushing, milling, processing, drilling, grinding, polishing and cutting materials containing quartz. Silicosis remains a public-health problem in emerging economies.

Regulations and strategies for controlling exposure have helped to reduce the incidence of silicosis in high-income countries. However, outbreaks among workers fabricating countertops from natural stone powders in resin binders demonstrate an unacceptable ignorance of this health hazard (*Lancet Respir. Med.* **7**, 283; 2019).

Steven Ronsmans, Benoit Nemery *Centre for Environment and Health, Leuven, Belgium.*
steven.ronsmans@kuleuven.be

Sand: save it for sea-level rise

Mette Bendixen and colleagues point out that sand extracted from fluvial environments is being consumed faster than it is produced (*Nature* **571**, 29–31; 2019) This has deep implications for managing flood risk in a changing climate.

Extracting sand or restricting its movement (such as through river damming) reduces sediment availability. This means that when large floods occur, insufficient sediment is deposited on the land for it to act as a defence against smaller floods. Fluvial-sediment depletion can also lead to coastal erosion, especially if accompanied by illegal sand mining on the foreshore.

Sea-level rise is projected to accelerate in the second half of this century. According to Bendixen and colleagues, sand prices could be exceptionally high by then. Instead of squandering sand, we need to save it.

Sally Brown, Susan Hanson *University of Southampton, UK.*
sb20@soton.ac.uk

Rule out nepotism in psychology awards

The payment of substantive fees to some psychologists who give talks on their own research has sparked concerns over conflicts of interest (COIs; *Nature* **571**, 20–23; 2019). We cannot rule out the possibility that the handing out of academic awards and prizes in psychology by professional societies or associations might also be subject to COIs.

We scrutinized the websites of 58 psychology societies using a pre-registered protocol (A. H. Stoevenbelt *et al.* Preprint at https://psyarxiv.com/phyu3; 2019). Our aim was to determine whether we could exclude the possibility that any recipients of such awards were closely affiliated with individuals on the award committees — for example, as family members, collaborators, mentees or colleagues.

Most of the societies (72.4%) failed to highlight any potential COIs in the committees responsible for selecting award winners. Less than half of them (44.8%) published no COI regulations at all. And, of those that did, only half (27.6%) explicitly mentioned avoiding COIs in choosing prizewinners.

We urge psychology societies to avoid conveying the impression of hidden nepotism by openly publishing their policies on personal COIs.

Andrea H. Stoevenbelt *Tilburg University, the Netherlands.*
a.h.stoevenbelt@uvt.nl
**On behalf of 4 correspondents; see go.nature.com/2zj9y5k.*

CORRECTION

In the Nature Index 2019 Annual Tables (*Nature* **570**, S1–S6; 2019) the fractional counts, percentage changes and article counts used for the tables were incorrect, which affected the rankings of some institutions. The updated data, graphics and rankings can be found online at https://www.nature.com/collections/fbfjafhcbb.

TUMOUR BIOLOGY

Cancer-cell death ironed out

Ferroptosis is a form of cell death. The finding that cells that have certain mutations in the Hippo signalling pathway are susceptible to ferroptosis might offer a way to treat a cancer called mesothelioma. SEE LETTER P.402

DEAN FENNELL

In the late twentieth century, there was a rise in a type of cancer called mesothelioma, which is caused by exposure to asbestos used in building materials. Mesothelioma often arises decades after exposure, accounting for tens of thousands of deaths annually worldwide[1]. Even with the treatments currently available, it is inevitably fatal. There is therefore an urgent need to develop more effective therapies for this type of cancer. Wu *et al.*[2] report on page 402 that mutations in a cell-signalling pathway that commonly occur in mesothelioma create a tumour vulnerability that might be targeted to treat this disease.

Mesothelioma most often originates in the lining of the lungs, in cells that form the pleural membrane. Mutations frequently found in mesothelioma cells often inactivate proteins, called tumour suppressors, that function in anticancer pathways. One of the most common such inactivated proteins is called merlin (encoded by the *NF2* gene), which functions in the highly evolutionarily conserved Hippo signalling pathway. This pathway was originally identified in the fruit fly *Drosophila melanogaster*[3,4], and it comprises a signalling cascade that controls cell proliferation and organ size. If merlin or another protein in this pathway, such as LATS2, is inactivated, downstream proteins called YAP and TAZ can boost the expression of genes that promote tumour formation. Certain cancers can even become 'addicted' to YAP-mediated transcription for their survival[5].

However, if merlin, LATS2 and another protein called LATS1 are functional, YAP and TAZ undergo phosphorylation (a phosphate group is attached to them), which modifies the proteins and blocks their function by preventing them from entering the nucleus to drive gene expression[6]. Mutations in the genes encoding merlin and LATS2 are positively selected during tumour development[7], consistent with their normal roles as tumour-suppressor proteins in mesothelioma.

Wu and colleagues studied the gene-expression profiles of human cancer cells grown *in vitro*, and report that YAP and TAZ drive the expression of proteins, such as ACSL4, that are needed for a type of cell death called ferroptosis. The authors also uncovered a connection between the ability of cells to suppress ferroptosis and the cell–cell contact that depends on the protein E-cadherin. The authors report that high expression of E-cadherin in human mesothelioma cells grown *in vitro* is associated with resistance to ferroptosis. E-cadherin activates the Hippo pathway, and the authors went on to explore the relationship between this pathway and ferroptosis.

Cell death that occurs through ferroptosis depends on a reaction between cellular iron and hydrogen peroxide[8]. During ferroptosis, a polyunsaturated fatty acid — a type of lipid found in the cell membrane — undergoes a modification called peroxidation, which causes an increase in the level of molecules termed reactive oxygen species. Ferroptosis is often linked to depletion of the amino acid cysteine, which is imported into cells by the protein SLC7A11. Cysteine provides a building block for the production of glutathione, a molecule involved in a pathway that can combat ferroptosis.

The drug sorafenib is approved for clinical use. It can induce ferroptosis by inhibiting SLC7A11. The authors demonstrate that sorafenib treatment of cultured human mesothelioma cells that have mutations in the gene encoding merlin causes the cells to undergo ferroptosis. They report that this sensitivity to ferroptosis depends on YAP- and TAZ-mediated gene expression (Fig. 1).

Two independent clinical trials[9,10] found that sorafenib caused tumour shrinkage or stabilization in people with mesothelioma. However, neither trial evaluated the mutations present in the patients' tumours, and it is tempting to speculate that the tumours of people who responded particularly well had mutations that inhibited the Hippo signalling pathway and that thereby boosted

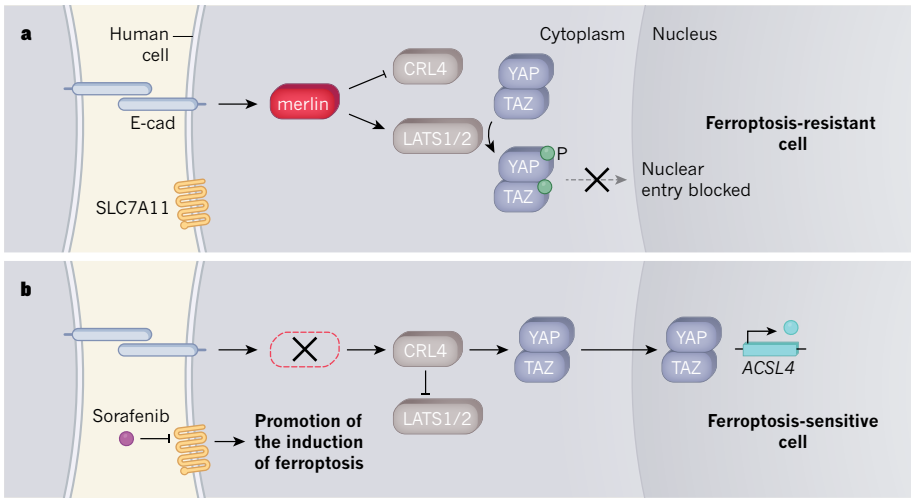


Figure 1 | Regulation of ferroptosis in human cells. Ferroptosis is a type of cell death whose induction is affected by a pathway that depends on the protein SLC7A11. Wu *et al.*[2] investigated how an anticancer signalling pathway called the Hippo pathway, in which mutations commonly occur in cancer cells, affects ferroptosis. **a**, Interactions between receptor proteins called E-cadherin (E-cad) on adjacent cells can trigger the Hippo pathway. A protein called merlin in this pathway prevents cancer-promoting gene expression by inhibiting a protein called CRL4. CRL4 inhibition enables the proteins LATS1 and LATS2 to add a phosphate group (P) to the proteins YAP and TAZ, and this phosphorylation prevents the proteins from entering the nucleus and driving gene expression. The authors report that YAP and TAZ drive the expression of genes that promote ferroptosis, revealing that Hippo pathway signalling makes cells resistant to ferroptosis. **b**, If merlin is not expressed because of a mutation, CRL4 is not inhibited and LATS1 and LATS2 cannot function. YAP and TAZ can enter the nucleus and drive the expression of genes, such as *ACSL4*, that promote ferroptosis. The authors report that tumour cells that lack merlin can undergo ferroptosis if treated with an inhibitor of SLC7A11, called sorafenib.

YAP- and TAZ-mediated gene expression.

Might other mutations beyond those in the Hippo pathway also regulate ferroptosis in mesothelioma? The most commonly mutated gene in this cancer[11] encodes the tumour-suppressor protein BAP1. This enzyme affects gene expression, and can cause a reduction in the expression of SLC7A11, which, in turn, leads to ferroptosis[12]. If the gene that encodes BAP1 is mutated, ferroptosis does not occur[12]. Therefore, the presence of wild-type BAP1 might help to enhance ferroptosis, along with any boost to ferroptosis provided by the use of SLC7A11 inhibitors. It is not known whether drugs that induce ferroptosis, such as sorafenib, would be effective in cells in which mutations inactivate BAP1.

Other approaches to targeting mesothelioma in which the Hippo pathway is inactivated are being explored. For example, in animal studies, loss of merlin expression is associated with cancer-cell vulnerability to inhibition of a protein called focal adhesion kinase[13]. However, no clinical benefit was found with this approach in a clinical trial[14]. Direct targeting of the interaction between YAP and TEAD, a protein to which YAP binds when it drives gene expression, is another strategy being pursued to block cancer-promoting gene expression[15]. Finally, YAP and TAZ recruit the protein BRD4 to drive the expression of specific genes, and use of a small-molecule inhibitor to target BRD4 can disrupt YAP- and TAZ-mediated gene expression[16]. This class of small-molecule inhibitor is entering early clinical trials. All of these approaches aim to block YAP- and TAZ-mediated gene expression. However, if the anticancer strategy being used aimed to trigger ferroptosis in mesothelioma cells, then YAP- and TAZ-mediated gene expression would be required.

Identifying a tumour that has an inactivated Hippo signalling pathway as a means of a developing personalized cancer therapy — the ultimate goal — poses some challenges for mesothelioma. Focusing only on tumours that have lost merlin function would probably miss mesotheliomas in which Hippo signalling is inhibited by inactivation of other proteins, such as LATS1 and LATS2. A previous study[17] of the Hippo pathway in various cancers has revealed that 22 genes are commonly transcribed by YAP and TAZ, and this transcriptional profile might offer a way to identify ferroptosis-sensitive tumours. Furthermore, because this profile was found[17] in several types of tumour, triggering ferroptosis might be worth exploring for cancers other than mesothelioma.

Wu and colleagues' report highlights a strategy that could offer a way of developing a personally tailored anticancer therapy. However, therapies targeted to mutations in an individual's mesothelioma are still in their infancy. Clinical trials that take this approach, for example the mesothelioma stratified therapy trial in which I am involved (see go.nature.com/2o19lah), might help to make progress in such endeavours, and provide improved treatments at a time of unmet clinical need. ■

Dean Fennell *is at the Mesothelioma Research Programme, Leicester Cancer Research Centre, University of Leicester, and at the University Hospitals of Leicester NHS Trust, Leicester LE2 7LX, UK.*
e-mail: df132@leicester.ac.uk

1. Odgerel, C.-O. *et al. Occup. Environ. Med.* **74**, 851–858 (2017).
2. Wu, J. *et al. Nature* **572**, 402–406 (2019).
3. Wu, S., Huang, J., Dong, J. & Pan, D. *Cell* **114**, 445–456 (2003).
4. Udan, R. S., Kango-Singh, M., Nolo, R., Tao, C. & Halder, G. *Nature Cell Biol.* **5**, 914–920 (2003).
5. Han, H. *et al. Oncogene* **37**, 6414–6424 (2018).
6. Li, W. *et al. Cancer Cell* **26**, 48–60 (2014).
7. Martincorena, I. *et al. Cell* **171**, 1029–1041 (2017).
8. Dixon, S. J. *et al. Cell* **149**, 1060–1072 (2012).
9. Papa, S. *et al. J. Thorac. Oncol.* **8**, 783–787 (2013).
10. Dubey, S. *et al. J. Thorac. Oncol.* **5**, 1655–1661 (2010).
11. Hmeljak, J. *et al. Cancer Discov.* **8**, 1549–1565 (2018).
12. Zhang, Y. *et al. Nature Cell Biol.* **20**, 1181–1192 (2018).
13. Shapiro, I. M. *et al. Sci. Transl. Med.* **6**, 237ra68 (2014).
14. Fennell, D. A. *et al. J. Clin. Oncol.* **37**, 790–798 (2019).
15. Liu-Chittenden, Y. *et al. Genes Dev.* **26**, 1300–1305 (2012).
16. Zanconato, F. *et al. Nature Med.* **24**, 1599–1610 (2018).
17. Wang, Y. *et al. Cell Rep.* **25**, 1304–1317 (2018).

The author declares competing financial interests. See go.nature.com/32xpoox for details.

This article was published online on 24 July 2019.

PLANETARY SCIENCE

Signs that Jupiter was mixed by a giant impact

Simulations suggest that Jupiter's dilute core might be the result of a collision between the planet and a Uranus-mass planetary embryo. This finding indicates that giant impacts could be common during planet formation. SEE LETTER P.355

TRISTAN GUILLOT

In the past couple of years, NASA's Juno spacecraft has measured Jupiter's gravitational field with exquisite accuracy[1,2]. The results have revealed that the planet's fluid hydrogen–helium envelope does not have a uniform composition: the inner part contains more heavy elements than the outer part[3,4]. On page 355, Liu *et al.*[5] propose that this asymmetry resulted from a head-on collision between the young Jupiter and a planetary embryo that had a mass about ten times that of Earth. The authors suggest that the primordial cores of the planet and of the embryo would have merged and then partially mixed with Jupiter's envelope, explaining the structure of the planet seen today.

Scars of impacts abound on rocky planetary bodies. For example, the Moon is covered in craters, and was formed by a collision that occurred 4.5 billion years ago between Earth and a massive body[6]. Although impacts leave no direct imprint on the surfaces of fluid planets, the tilts of the rotational axes of Saturn (27°), Uranus (98°) and Neptune (30°) might indicate that violent collisions occurred in the past[7]. After all, it is known that massive planetary embryos on the order of ten Earth masses must have been present in the early Solar System[8], in addition to the planets that are still here. Jupiter, with its small tilt (3°), seems to have escaped unscathed[7]. But according to Liu and colleagues, this was not the case.

Jupiter is mostly made of hydrogen and helium. However, observations of its atmospheric composition[9] and gravitational field show that it contains a non-negligible proportion of heavier elements in the form of a central core and in the hydrogen–helium envelope. This envelope is fluid and is expected to be largely convective[10], so it was surprising when Juno revealed that the envelope's composition is not uniform. Instead, the core seems to be partially diluted in the envelope, extending to almost half of the planet's radius[3,4] (Fig. 1).

Producing this internal structure directly would require the delivery (accretion) of 10–20 Earth masses[3,4] of heavy elements to the young Jupiter after the core had formed and during the first half of the growth of the envelope. The accretion of this material would need to have stopped after the planet had grown to about half of its present mass.

Formation models indicate that this hypothesis is unlikely. In these models, when Jupiter reaches about 30 Earth masses, the growth of the envelope by accretion is fast[11], and the planet efficiently pushes away any dust particle that is millimetre-sized or larger[12]. As a result, the envelope should be poor in heavy elements. Any subsequent delivery of heavy elements by planetesimals (the asteroid-sized precursors of planets) or small planets is inefficient and cannot explain a heavy-element abundance that would increase with depth, as is observed. Erosion of the core into the envelope is possible[10,13], but simulations show that this process tends to remove any small composition gradients that exist in the

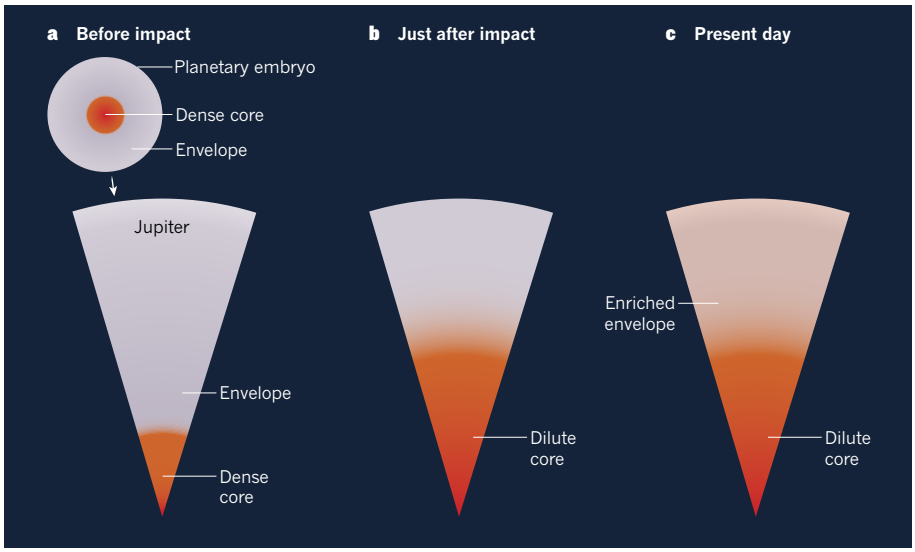


Figure 1 | Three phases of Jupiter. Liu *et al.*[5] propose that the present-day internal structure of Jupiter is the result of a giant impact between the young planet and a planetary embryo that had roughly the mass of Uranus. **a**, In the authors' model, before the impact, both Jupiter and the embryo contained a dense central core of heavy elements and a hydrogen–helium envelope. The colours represent the density of material, ranging from low (white) to high (dark orange). **b**, Just after the impact, the two cores merged and partially mixed with the planet's envelope to produce a dilute core. **c**, After subsequent evolution, the dilute core remained, but was partially eroded into the envelope, causing the envelope to be enriched in heavy elements.

envelope, rather than increase them[14].

The solution proposed by Liu *et al.* is simple. In their model, a planetary embryo that has a dense core of heavy elements collides with the forming Jupiter. The cores of the two bodies then merge and become partially mixed with Jupiter's envelope. This explanation requires a massive embryo (of about ten Earth masses) and an impact that is somewhat head-on, but these two requirements seem reasonably likely. The authors show that cooling and subsequent convective mixing of the outer part of the envelope mixes only some of the heavy elements, leaving the planet's dilute core relatively unaffected (Fig. 1). In one fell swoop, this picture might therefore explain the dilute core detected by Juno[3,4] and the global abundance of heavy elements in Jupiter's atmosphere[9].

Liu and colleagues' model should now be refined. In particular, it needs to be coupled to realistic scenarios for the formation of the Solar System[8]. Moreover, the mixing of heavy elements in the model should take into account heat and element diffusion — a process known as diffusive convection[13]. The results should also be compared quantitatively with constraints on Jupiter's gravitational field from Juno[1,2] and on the planet's atmospheric composition obtained from spectroscopy[10].

The authors' model indicates that giant impacts might frequently occur during planet formation. This possibility could account for the tilts of the planets in the Solar System. It might also explain how some giant exoplanets, known as hot Jupiters, have accreted more than 100 Earth masses of heavy elements[15,16] — a feature that is extremely difficult to obtain from conventional formation models. Hot Jupiters are situated close to their host stars, in regions in which the gravitational pull of the star is extremely strong. As a result, these exoplanets might be able to collect planetary embryos efficiently through a series of giant impacts, rather than ejecting them, and thus increase their heavy-element content.

Although giant planets have a fluid surface that cannot record traces of impact events, such planets hold clues to a violent past that led to the planetary systems observed today. The model proposed by Liu *et al.* enables present-day observations to be linked to the early days of the formation of the Solar System. Progress will come from an extension of studies such as this one to giant planets around the Sun and other stars. A continued exploration of the Solar System is crucial, particularly of Uranus and Neptune, which might be thought of as leftovers from a large population of massive planetary embryos in the early Solar System. ■

Tristan Guillot *is in the Université Côte d'Azur, Laboratoire Lagrange, Observatoire de la Côte d'Azur, CNRS UMR 7293, 06304 Nice Cedex 4, France.*
e-mail: tristan.guillot@oca.eu

1. Folkner, W. M. *et al. Geophys. Res. Lett.* **44**, 4694–4700 (2017).
2. Iess, L. *et al. Nature* **555**, 220–222 (2018).
3. Wahl, S. M. *et al. Geophys. Res. Lett.* **44**, 4649–4659 (2017).
4. Debras, F. & Chabrier, G. *Astrophys. J.* **872**, 100 (2019).
5. Liu, S.-F. *et al. Nature* **572**, 355–357 (2019).
6. Hartmann, W. & Davis, D. *Icarus* **24**, 504–515 (1975).
7. Chambers, J. & Mitton, J. in *From Dust to Life: The Origin and Evolution of Our Solar System* 216 (Princeton Univ. Press, 2017).
8. Izidoro, A., Morbidelli, A., Raymond, S. N., Hersant, F. & Pierens, A. *Astron. Astrophys.* **582**, A99 (2015).
9. Wong, M. H., Mahaffy, P. R., Atreya, S. K.,

Niemann, H. B. & Owen, T. C. *Icarus* **171**, 153–170 (2004).
10. Guillot, T., Stevenson, D. J., Hubbard, W. B. & Saumon, D. in *Jupiter: The Planet, Satellites and Magnetosphere* (eds Bagenal, F., Dowling, T. E. & McKinnon, W. B.) 35–57 (Cambridge Univ. Press, 2004).
11. Mordasini, C., Alibert, Y., Klahr, H. & Henning, T. *Astron. Astrophys.* **547**, A111 (2012).
12. Paardekooper, S.-J. & Mellema, G. *Astron. Astrophys.* **425**, L9–L12 (2004).
13. Moll, R., Garaud, P., Mankovich, C. & Fortney, J. J. *Astrophys. J.* **849**, 24 (2017).
14. Vazan, A., Helled, R. & Guillot, T. *Astron. Astrophys.* **610**, L14 (2018).
15. Moutou, C. *et al*. *Icarus* **226**, 1625–1634 (2013).
16. Thorngren, D. P., Fortney, J. J., Murray-Clay, R. A. & Lopez, E. D. *Astrophys. J.* **831**, 64 (2016).

MICROBIOLOGY

No bacteria found in healthy placentas

Analysis of hundreds of placentas provides convincing evidence that this organ does not harbour microorganisms that can enter the fetal gut — a key finding for research into how the human microbiota is established. SEE ARTICLE P.329

NICOLA SEGATA

The early human embryo is free of microorganisms, whereas the post-weaning infant hosts a community of microbes — a microbiota — comparable in complexity to that in adults. How and when the symbiosis between a human and their microbiota is established are subjects of active research. On page 329, de Goffau *et al.*[1] provide evidence that the placenta, which acts as the interface between the maternal body and the fetus, is not colonized by microorganisms in healthy pregnancies and is thus unlikely to be the main gateway for the development of the infant microbiota *in utero*.

If the microbial colonization of humans occurs in the womb, then this would have key implications for the shaping of the early immune system. An infant's first stool is already populated with microorganisms, but it is unclear whether this is solely the result of microbial acquisition during[2] and after[3] delivery, or if microbes also reach and colonize the fetus before birth. Because sampling fetal gut content is much more difficult than collecting the placenta and amniotic fluid during (elective) caesarean delivery, scientists have focused on the latter two at the interface between the maternal and fetal bodies. The conclusive identification of microbial communities in and on the placenta would indeed suggest that microbes colonize the fetus, but, in the past few years, evidence has been presented both that supports[4–7] and that refutes[8–11] the long-standing dogma that the placenta and amniotic fluid are sterile in physiological conditions — that is, during healthy pregnancy. The debate about this issue therefore remains open[12,13] (Fig. 1).

It is not disputed that, during a healthy pregnancy, the placenta and amniotic fluid cannot host a concentration of bacteria as high as that observed in the adult mouth or gut. The technical challenge in studies of placenta samples is therefore to distinguish any microorganisms that are truly present in small quantities on these tissues from those found on laboratory tools and from contamination of the samples during collection. Small amounts of microbial contamination can be pervasive, and sources range from the air to supposedly sterile DNA-extraction kits[14] and other items associated with DNA processing and sequencing[15]. There was thus a need for studies to rigorously account for potential contamination; these studies would also need a sufficiently large sample size to ensure statistical robustness. De Goffau and colleagues now report on such a study.

The authors analysed placenta samples from 537 women — by far the largest number of samples used in a study of this kind — using a thorough DNA-sequencing approach to search for microbial content. They used the same DNA-extraction toolkit and sequencing procedures on negative controls — 'blank' samples that were supposedly free from biological material. They also used positive controls, produced by spiking placental samples with a known amount of the bacterium *Salmonella bongori*, to calibrate the abundance of other microbes that might be in the sample. The sequencing was performed using two complementary techniques, known as shotgun metagenomics[16] and 16S rRNA gene amplicon sequencing[17], to account for technique-specific potential biases. The results were clear: the placenta does not harbour microbes during healthy pregnancy, and contamination issues were a convincing explanation for the presence of any detected bacteria.

Some of the details reported in the paper reveal how pervasive contaminating microbes can be when concentrations of bacteria in the samples are very low. For example, two potential

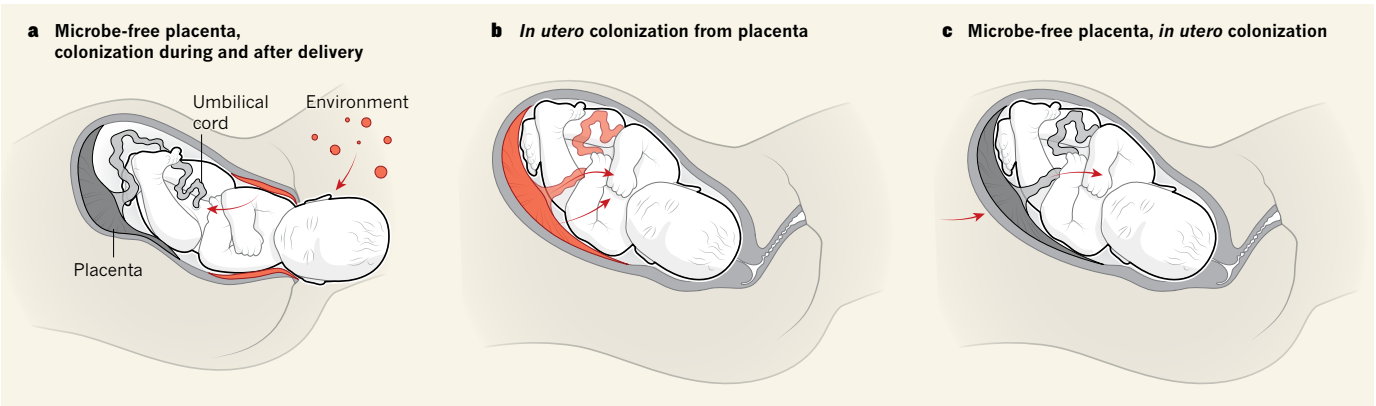


Figure 1 | Scenarios for bacterial colonization of the infant gut. a, It has long been thought that the human placenta and the fetus are free of microorganisms. Newborns were therefore expected to acquire gut bacteria from the mother during delivery and from the environment (red regions indicate sources of bacteria), with further influences associated with the mode of delivery and feeding regime (breastfeeding or formula milk). **b**, However, in the past few years, evidence has been published[8–11] suggesting that the placenta contains bacteria and that bacterial colonization of the fetal gut therefore occurs in the womb. **c**, *In utero* colonization of the fetal gut from the mother might also occur under certain circumstances, even if the placenta is microbe-free. De Goffau *et al.*[1] now report convincing evidence that the placenta is free of bacteria during healthy pregnancies, thus ruling out the scenario in **b**.

disease-causing bacteria,*Vibrio cholerae* and *Streptococcus pneumoniae*, were detected by shotgun metagenomics and matched strains of bacteria that had previously been sequenced on the same apparatus. The detection of these bacteria is therefore most probably the result of cross-contamination of the authors' sequencing machine. The ability of modern sequencing methods to detect low numbers of bacteria is thus a problem in some experiments, because even tiny levels of contaminants can result in a false-positive detection. Greater contamination of the authors' samples occurred during the earlier stages of sample preparation than in later stages. The authors confirmed previous reports[14] stating that a relatively rich microbiota was present in commercial DNA-extraction kits, and identified company-specific communities of bacteria from the genetic material extracted from the blank control samples.

Overall, the complex procedures used by de Goffau and colleagues to identify contaminants allowed them to reach a clear conclusion: only one type of bacterium was convincingly found in the placental samples in their study, and it was in only about 5% of those samples. This finding provides strong evidence that there is no functional microbiota in the placenta and suggests that it is highly unlikely that infants acquire microbes from the placenta in normal physiological conditions.

The bacterium occasionally detected in the placenta was *Streptococcus agalactiae*. If present in the mother during childbirth, *S. agalactiae* can be transmitted to the newborn and cause pneumonia, septicaemia and meningitis; several clinical practices are used to prevent such transmission[18]. The identification of *S. agalactiae* in some of the placenta samples in the study does not conflict with the dogma that the womb is microbe-free in healthy pregnancies, because this bacterium is associated with disease. Indeed, the finding that *S. agalactiae* is the only bacterium to be found on the placenta, and in a low number of samples, mirrors the expectation that a small fraction of pregnant mothers are infected with it, and that it can undergo intrauterine transmission — therefore adding credibility to the experimental findings.

De Goffau and colleagues' carefully controlled, large-scale study was needed to provide strong evidence for the absence of bacteria in the placenta. As such, the study also sets a benchmark for investigations dealing with other human organs or tissues that, at most, carry a small number of bacteria, such as the lungs or blood. Nevertheless, negative results are hard to prove conclusively, so the dogma that the womb is free of microbes should be further investigated. Bacteria can overcome many host barriers under certain conditions, and just one bacterial cell that reaches the gut of the fetus could potentially start *in utero* colonization. How the symbiosis of a human host with their microbiota is established remains an intriguing, fundamental question, but we can now be confident that the placenta is not a microbial reservoir and therefore is not a major direct stream of diverse microbes to the fetus under healthy conditions. ■

Nicola Segata *is in Department CIBIO, University of Trento, Trento 38123, Italy.*
e-mail: nicola.segata@unitn.it

1. de Goffau, M. C. *et al. Nature* **572**, 329–334 (2019).
2. Ferretti, P. *et al. Cell Host Microbe* **24**, 133–145 (2018).
3. Korpela, K. *et al. Genome Res.* **28**, 561–568 (2018).
4. Leon, L. J. *et al. Appl. Environ. Microbiol.* **84**, e00483-18 (2018).
5. Leiby, J. S. *et al. Microbiome* **6**, 196 (2018).
6. Theis, K. R. *et al. Am. J. Obstet. Gynecol.* **220**, 267–267 (2019).
7. Lim, E. S., Rodriguez, C. & Holtz, L. R. *Microbiome* **6**, 87 (2018).
8. Collado, M. C., Rautava, S., Aakko, J., Isolauri, E. & Salminen, S. *Sci. Rep.* **6**, 23129 (2016).
9. Urushiyama, D. *et al. Sci. Rep.* **7**, 12171 (2017).
10. Aagaard, K. *et al. Sci. Transl. Med.* **6**, 237ra65 (2014).
11. Antony, K. M. *et al. Am. J. Obstet. Gynecol.* **212**, 653 (2015).
12. Perez-Muñoz, M. E., Arrieta, M. C., Ramer-Tait, A. E. & Walter, J. *Microbiome* **5**, 48 (2017).
13. Bushman, F. D. *Am. J. Obstet. Gynecol.* **220**, 213–214 (2019).
14. Salter, S. J. *et al. BMC Biol.* **12**, 87 (2014).
15. Eisenhofer, R. *et al. Trends Microbiol.* **27**, 105–117 (2019).
16. Quince, C., Walker, A. W., Simpson, J. T., Loman, N. J. & Segata, N. *Nature Biotechnol.* **35**, 833–844 (2017).
17. Hamady, M. & Knight, R. *Genome Res.* **19**, 1141–1152 (2009).
18. Johri, A. K. *et al. Nature Rev. Microbiol.* **4**, 932–942 (2006).

This article was published online on 31 July 2019.

CELL BIOLOGY

How plants perceive salt

High salt levels in the soil harm plant growth and limit crop yields. A salt–binding membrane lipid has been identified as being essential for salt perception and for triggering calcium signals that lead to salt tolerance. SEE ARTICLE P.341

LEONIE STEINHORST & JÖRG KUDLA

Salt as a nutrient for humans is a double-edged sword, being tasty in small amounts but generating an adverse response as the concentration rises. Distinct protein receptors have been shown to mediate these opposing reactions in animals. Excessive uptake of salt is not only unhealthy for humans but also detrimental for plants, because high levels of salt in the soil limit plant growth and crop yields. This is of concern, given that such conditions affect approximately 7% of land globally, including areas used for agriculture, and high salinity affects about 30% of irrigated crops[1]. On page 341, Jiang *et al.*[2] shed light on how plants recognize salt in their surroundings.

The salt sodium chloride (NaCl) is the main cause of salt stress in plants. It is toxic to cells because at high intracellular concentrations, Na^+ ions compete with other ions for involvement in biological reactions. It also has a negative effect on cellular functions by perturbing the balance of ions and thus of water — generating what is called an osmotic perturbation. It was not known how plants perceive stress generated by high salt and whether they can distinguish between ionic and osmotic perturbations.

The exposure of plants to salt stress triggers an immediate temporally and spatially defined rise in the concentration of cytoplasmic calcium ions (Ca^{2+}). It is thought that a calcium channel, of as yet unknown identity, provides a route for Ca^{2+} to enter cells during such calcium signalling. This Ca^{2+} signal leads to cellular adaption to salt stress in plant roots, and the subsequent formation of Ca^{2+} waves that spread over long distances and mediate adaptation responses throughout the entire plant[3,4]. Central to salt tolerance is the evolutionarily conserved SOS pathway. In this pathway, proteins such as SOS3, which can bind Ca^{2+} ions, decode the Ca^{2+} signal and activate[5] a protein kinase enzyme called SOS2. This enzyme, in turn, activates a protein in the cell membrane called SOS1, which is a type of protein known as an antiporter that can transport Na^+ ions out of the cell. SOS2 also promotes the sequestration of Na^+ from the cytoplasm into an organelle called a vacuole[6]. However, the components and mechanisms governing the perception of extracellular Na^+ and driving salt-induced Ca^{2+} signalling were unknown.

"It was not known how plants perceive stress generated by high salt."

Jiang and colleagues performed a genetic screen using the model plant *Arabidopsis thaliana* to identify mutant plants that had an abnormally low Ca^{2+}-signalling response to high Na^+ exposure, but that could still generate Ca^{2+} signals when challenged with other types of stress. Taking this approach, they identified a plant that had a mutation

in the gene encoding the protein IPUT1. IPUT1 acts at a central step required for the synthesis of a type of lipid called a sphingolipid. This is surprising because, in animals, Na^+ ions are sensed by protein receptors rather than through the involvement of lipids.

IPUT1 catalyses the formation of the lipid glycosyl inositol phosphorylceramide (GIPC). GIPCs are major constituents of the outer layer of the lipid bilayer in the plasma membranes of plants, accounting for up to 40% of plasma-membrane lipids, and they can be considered equivalent in function to lipids called sphingomyelins that are found in animals[7].

Other mutations previously identified[8] in the gene for IPUT1 severely affect plant development; the mutation studied by the authors did not impair development, however, which enabled the role of this protein in the response to salt to be investigated. Emphasizing the importance of Ca^{2+} signalling for plant tolerance to high salt levels, the authors report that the abnormal Ca^{2+} signals and long-distance Ca^{2+} waves in these mutant plants were associated with the plants' high sensitivity to salt stress. Remarkably, these mutants showed no alterations in their resilience to comparably severe osmotic stress that was induced experimentally in ways that did not require the manipulation of Na^+ levels.

Jiang and colleagues report that salt-stress-triggered changes in membrane polarization (the difference in electrical charges between the interior and exterior of the cell) and activation of the SOS pathway were impaired in the mutant plants, compared with wild-type plants. The authors carried out biochemical tests revealing that GIPCs can bind Na^+ ions and other ions that have a single positive charge, such as potassium (K^+) and lithium (Li^+). This observation is interesting because there is evidence for an inverse relationship between the concentrations of K^+ and Na^+ in plant cells during salt stress[5]. It would be worth investigating whether and, if so, how K^+ binding GIPCs modulates the ability of GIPC to bind Na^+, and vice versa. Taken together, the authors' evidence supports their conclusion that direct binding of Na^+ by GIPCs is an essential step in sodium sensing in plants that then triggers the calcium signals that lead to salt-tolerance responses.

The authors propose that plant GIPCs function in the same way as a type of lipid called a ganglioside that is found in animal cells. In neuronal cells, gangliosides directly or indirectly regulate important properties of receptors and ion channels in specific regions of the plasma membrane known as microdomains, which have a distinctive lipid composition[9]. The authors suggest that, like ganglioside function in animals, GIPCs in plants interact directly with Ca^{2+} channels. Na^+ binding to GIPCs might modulate channel activity, leading to the generation of Ca^{2+} signals in the cell (Fig. 1a).

However, the evidence currently available also supports a different model, in which GIPCs stimulate Ca^{2+} signals through an indirect and more complex mechanism (Fig. 1b). There is growing evidence that microdomains in lipid membranes, and specifically GIPCs in these microdomains, aid the regulation of signalling in plants.

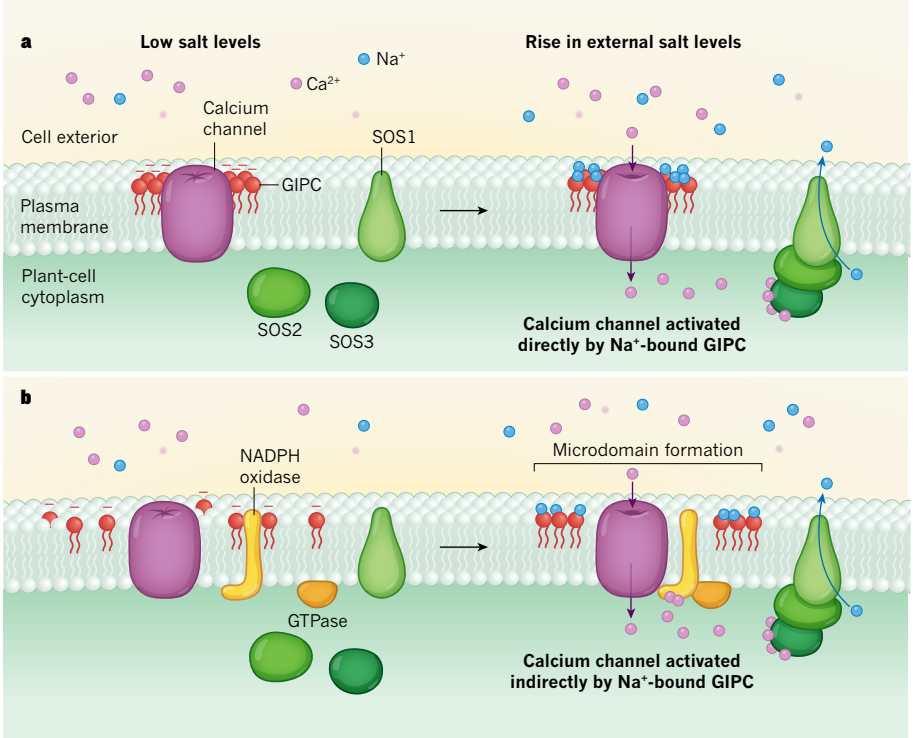


Figure 1 | How plants sense salt and activate calcium channels. **a**, When the sodium ions (Na^+) of salt are sensed outside a plant cell, an unknown calcium channel is activated and calcium ions (Ca^{2+}) enter the cell. Jiang *et al.*[2] reveal that a type of negatively charged membrane lipid called glycosyl inositol phosphorylceramide (GIPC) directly binds external Na^+ ions. The authors propose that a direct interaction between sodium-bound GIPC and the calcium channel leads to channel activation. The subsequent influx of Ca^{2+}drives an adaptive response to high salt levels in which the Ca^2-binding protein SOS3 activates the protein SOS2, which, in turn, activates the protein SOS1 to pump Na^+ out of the cell. **b**, An alternative model for the calcium-channel activation is that Na^+ binding to GIPCs drives the formation of a microdomain — a region of distinctive lipid composition — in the plasma membrane. This microdomain would alter the dynamics of signalling proteins (such as NADPH oxidases or GTPases) in the microdomain, which can affect Ca^{2+} signalling. By an unknown mechanism, Na^+ binding to GIPCs might alter the assembly and activity of proteins in the microdomain, indirectly activating the calcium channel.

Salt stress also triggers the generation of molecules called reactive oxygen species (ROS)[4,10], which can induce Ca^{2+} signalling in plants[11]. Moreover, salt stress affects the formation and dynamics of microdomains in the plasma membrane, consequently affecting the activity and lateral mobility (the speed and range of movements) of enzymes called NADPH oxidases that act in the production of ROS signals[12]. Such stress also affects the lateral mobility of enzymes called GTPases that regulate NADPH oxidases[12]. These changes in microdomain arrangement in response to salt stress depend on the GIPC composition of the plasma membrane[12,13].

It is therefore tempting to speculate that the binding of Na^+ ions or other positively charged ions to GIPCs modulates the dynamics and assembly of protein complexes in microdomains. Thus, Na^+ binding to GIPCs might lead to the assembly of signalling complexes in a microdomain that enables a Ca^{2+} signal to be generated in response to salt-induced stress. In this way, Ca^{2+}-ion-channel activation might be an indirect consequence of Na^+ binding to GIPCs, and might involve the dynamic assembly and activation of other signalling proteins (such as NADPH oxidases) in these microdomains. It would be interesting to investigate whether SOS1 might be incorporated into such a microdomain.

There is evidence in plants that another type of membrane lipid called phosphatidylserine can also affect the formation of microdomains that mediate the regulation of GTPases, Ca^{2+} or ROS signalling[13]. It has been reported[14] that phosphatidylserine can regulate GTPase-mediated signalling in plants and enable the formation of hormone-induced (rather than salt-stress mediated) clustering of GTPases in lipid membranes. Moreover, GIPCs can contribute to the generation of other signalling events in plants. For example, they act as receptors for specific toxins that cause plant disease, and plants with altered GIPC composition are more resistant to such toxins than are plants with a normal GIPC composition[15]. These observations, together with those reported

50 Years Ago

With the growth of telecommunications based on geostationary orbits, there is growing concern that satellites may become so closely crowded together that they interfere with each other … An article in the current issue of the *Proceedings of the Institution of Electrical Engineers* … consists of a calculation of the capacity of the equatorial orbit to accumulate geostationary communications satellites. Their chief conclusion is that the capacity of the equatorial orbit, with present arrangements, is probably limited to about 2,000 telephone circuits for each degree of the orbit. For practical purposes, this amounts to roughly one satellite in each four degrees of the orbit, which in turn implies that it may take very little further development before parts of the equatorial orbit — over the Atlantic and America, for example — may be overcrowded.
From *Nature* 16 August 1969

100 Years Ago

The war has been responsible for great developments in many branches of science … [C]lose attention has been given to the subject of marine physics … especially … submarine acoustics … The singular property which distinguishes a submarine from other ships is its capacity of rendering itself invisible when pursued or when seeking and attacking its prey. Robbed of this power, it is an extremely vulnerable craft … The acoustic method of detecting a submerged submarine … was found to be far more sensitive and to give a much longer range than all other methods. Instruments used for this purpose are called hydrophones. … [T]he improved hydrophones developed for war service should greatly reduce the dangers of collisions and shipwreck.
From *Nature* 14 August 1919

by Jiang and colleagues, indicate that GIPCs fulfil versatile sensing and signalling functions in plants. This work also points to a crucial role for membrane-lipid composition in organizing functionally important signalling domains for many key processes in plants. ■

Leonie Steinhorst *and* **Jörg Kudla** *are at the Institute of Plant Biology and Biotechnology, University of Münster, Münster 48149, Germany.*
e-mails: l_stei02@uni-muenster.de; jkudla@uni-muenster.de

1. Schroeder, J. I. *et al. Nature* **497**, 60–66 (2013).
2. Jiang, Z. *et al. Nature* **572**, 341–346 (2019).
3. Choi, W. G., Toyota, M., Kim, S. H., Hilleary, R. & Gilroy, S. *Proc. Natl Acad. Sci. USA* **111**, 6497–6502 (2014).
4. Steinhorst, L. & Kudla, J. *Curr. Opin. Plant Biol.* **22**, 14–21 (2014).
5. Yang, Y. & Guo, Y. *New Phytol.* **217**, 523–539 (2018).
6. Kim, B. G. *et al. Plant J.* **52**, 473–484 (2007).
7. Cacas, J. L. *et al. Plant Physiol.* **170**, 367–384 (2016).
8. Tartaglio, V. *et al. Plant J.* **89**, 278–290 (2017).
9. Posse de Chaves, E. & Sipione, S. *FEBS Lett.* **584**, 1748–1759 (2010).
10. Manishankar, P., Wang, N., Köster, P., Alatar, A. A. & Kudla, J. *J. Exp. Bot.* **69**, 4215–4226 (2018).
11. Pei, Z. M. *et al. Nature* **406**, 731–734 (2000).
12. Nagano, M. *et al. Plant Cell* **28**, 1966–1983 (2016).
13. Hao, H. *et al. Plant Cell* **26**, 1729–1745 (2014).
14. Platre, M. P. *et al. Science* **364**, 57–62 (2019).
15. Lenarčič, T. *et al. Science* **358**, 1431–1434 (2017).

This article was published online on 31 July 2019.

ASTROPHYSICS

X marks the spot for fast radio bursts

Fast radio bursts are enigmatic astronomical signals that originate from deep in extragalactic space. Observations using an array of radio telescopes have identified a likely host galaxy for one of these signals. SEE LETTER P.352

JASON HESSELS

In 2007, astronomers detected a flash of radio waves that was much shorter in duration than the blink of an eye[1]. Such signals, now called fast radio bursts (FRBs), are thought to have been produced billions of years ago in distant galaxies[2]. If so, the sources of FRBs must be spectacularly energetic and, quite possibly, unlike anything that has ever been observed in our Galaxy. Pinpointing the galaxies that host FRBs is the key to unlocking the mysterious origins of these signals. On page 352, Ravi *et al.*[3] report the discovery of the likely host galaxy of an FRB that travelled for 6 billion years before reaching Earth. The properties of this galaxy suggest that active star formation is not essential for making an FRB source.

The maxim 'location, location, location' applies to FRBs: knowing where these signals originate is crucial to understanding what generates them. Although astronomers have detected almost 100 FRB sources so far[2], the measured positions of these sources on the sky have typically been too inaccurate to identify their host galaxies. One exception is the first FRB source observed to produce repeat bursts[4]. This source was localized to a region of active star formation in a puny 'dwarf' galaxy[5]. The finding supported theories that ascribe the origin of FRBs to the extremely condensed remnants of powerful stellar explosions called supernovae. For example, the repeating FRBs could originate from young and hyper-magnetized neutron stars — the collapsed remnants of massive stars[6].

However, most FRB sources have not been seen to produce repeat bursts. Astronomers have therefore questioned whether these apparently one-off events have a different origin from that of the repeating FRBs[2]. From a practical point of view, one-off FRBs are much more challenging to study than repeaters. In the case of a repeating FRB, a patient observer can wait for further bursts and refine the measured position of the source. But for a one-off FRB, the position needs to be pinpointed by capturing the necessary high-resolution data at the same time as the burst is discovered.

Ravi and colleagues achieved this feat using an array of ten relatively small (4.5-metre-diameter) radio dishes spread across an area of roughly one square kilometre in Owens Valley, California. This distributed telescope network, known as the Deep Synoptic Array 10-antenna prototype (DSA-10), can scan a broad swathe of sky for FRBs (Fig. 1a). It can also provide enough spatial resolution to determine the position of a burst on the sky with high precision[7]. This precision must indeed be extremely high: unless the position is known to 1,000th of a degree, robustly associating an FRB with a specific host galaxy is impossible[8]. Even though Ravi *et al.* determined the position of their FRB to this level of precision (Fig. 1b), there is still some uncertainty as to whether or not the identified galaxy is the true host.

The authors demonstrate that this likely

IMAGES: CALTECH/OVRO/V. RAVI

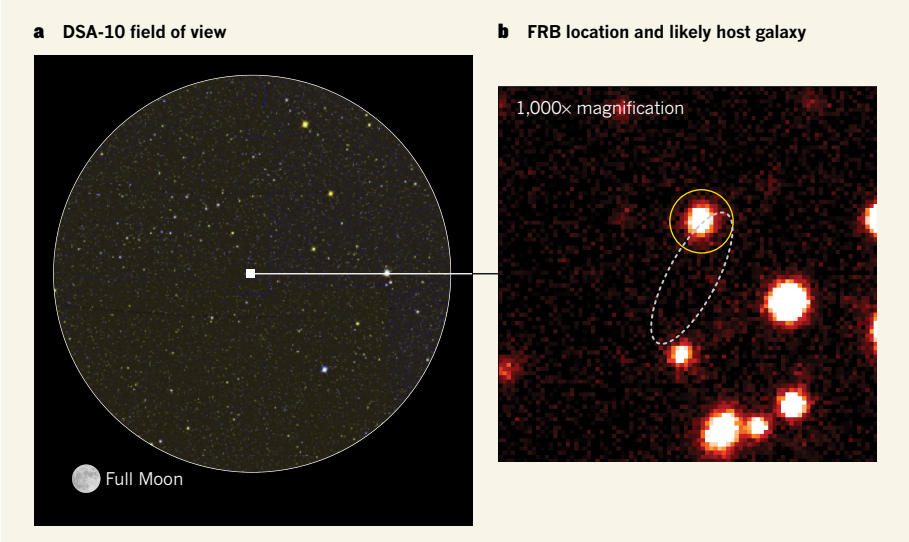


Figure 1 | Localization of a fast radio burst (FRB). **a**, Ravi *et al.*[3] report observations from an array of radio telescopes known as the Deep Synoptic Array 10-antenna prototype (DSA-10). The field of view of DSA-10 is roughly 40 square degrees[7], which is about 200 times the area on the sky that is covered by the full Moon when viewed from Earth's surface. **b**, Ravi and colleagues used DSA-10 to precisely determine the position of an FRB — a millisecond-duration flash of radio waves. The broken white ellipse shows the region in which the FRB could be located. The authors then identified a massive galaxy (indicated by the yellow circle) that is the likely host of the FRB.

host galaxy is markedly different from the host[5] of the well-localized source of the repeating FRB. It is 1,000 times more massive, and shows none of the prodigious star formation that is associated with the environment of the repeating-FRB source. Only a week before Ravi and colleagues' work was published online, a similar breakthrough was reported[9] using the Australian Square Kilometre Array Pathfinder (ASKAP) telescope. The authors of that paper achieved an even more precise localization of another one-off FRB, and also demonstrated that it originates from a massive galaxy that shows little signs of active star formation.

So, do these results mean that one-off FRBs and repeaters come from different galaxy types, and that they have physically different origins? Do astronomers have two puzzles on their hands? Perhaps, but with only three FRB host galaxies identified so far, many alternatives remain open. For instance, it is possible that all FRBs are generated by hyper-magnetized neutron stars, but that there are various ways in which such neutron stars can be produced[10]. Some might form directly through the collapse of a massive star, whereas others might be made from old neutron stars in a binary system that smash into each other as the orbital distance between them decreases. This difference could explain why some FRBs seem to originate from star-forming regions and others do not[10].

Excitingly, we will soon know a lot more. The mystery of FRBs has driven many teams worldwide to tune radio telescopes towards discovering and localizing these signals, and many thousands of FRBs are thought to happen somewhere on the sky each day[2]. The fact that fewer than 100 FRB sources have been detected is a reflection of the small fields of view of existing radio telescopes. If a sensitive radio telescope could be built that has a continuous view of the entire sky, FRBs would look like a fireworks display. However, wide-field telescopes such as the Canadian Hydrogen Intensity Mapping Experiment[11] (CHIME) are starting to change the game. It might not be long before astronomers have catalogued thousands of FRB sources and pinpointed at least dozens of them.

The precise localizations from DSA-10 and ASKAP are shedding light on the origins of FRBs, but they are also teaching us about the potential use of these signals as astronomical probes. FRBs are delayed in their arrival at Earth by the otherwise invisible material between galaxies[2]. By measuring the magnitude of this time delay, and comparing this measurement with the distance to the host galaxy, astronomers can map the density of ionized material in intergalactic space and thereby weigh the Universe in a unique way. The localizations of one-off FRBs suggest that FRB host galaxies will only slightly skew such measurements. Moreover, the results indicate that, with the detection and localization of thousands of FRBs, a 3D map of the material between galaxies could be made. ■

Jason Hessels *is at ASTRON (the Netherlands Institute for Radio Astronomy) and the Anton Pannekoek Institute for Astronomy, University of Amsterdam, 1098XH Amsterdam, the Netherlands.*
e-mail: j.w.t.hessels@uva.nl

1. Lorimer, D. R., Bailes, M., McLaughlin, M. A., Narkevic, D. J. & Crawford, F. *Science* **318**, 777–780 (2007).
2. Petroff, E., Hessels, J. W. T. & Lorimer, D. R. *Astron. Astrophys. Rev.* **27**, 4 (2019).
3. Ravi, V. *et al. Nature* **572**, 352–354 (2019).
4. Spitler, L. G. *et al. Nature* **531**, 202–205 (2016).
5. Bassa, C. G. *et al. Astrophys. J.* **843**, L8 (2017).
6. Margalit, B. & Metzger, B. D. *Astrophys. J.* **868**, L4 (2018).
7. Kocz, J. *et al.* Preprint at https://arxiv.org/abs/1906.08699 (2019).
8. Eftekhari, T. & Berger, E. *Astrophys. J.* **849**, 162 (2017).
9. Bannister, K. W. *et al. Science* https://doi.org/10.1126/science.aaw5903 (2019).
10. Margalit, B., Berger, E. & Metzger, B. D. Preprint at https://arxiv.org/abs/1907.00016 (2019).
11. The CHIME/FRB Collaboration. *Nature* **566**, 230–234 (2019).

TUMOUR BIOLOGY

A dynamic view of chemotherapy

Chemotherapy can halt cancer by causing cells to enter a non-dividing state called senescence, but sometimes it causes tumour cells to proliferate. It now seems that the dynamics of the protein p21 governs which of these fates occurs.

YUNPENG LIU & MICHAEL T. HEMANN

Chemotherapy usually works by inducing DNA damage that leads to cell death. However, rather than dying after chemotherapy, some tumour cells enter an inactive state, termed senescence, in which they are alive but have permanently stopped dividing[1]. Although senescence in normal cells drives ageing and tissue degeneration[2], cancer-therapy-induced senescence is associated with positive clinical outcomes[3]. Understanding the factors that drive the senescence of tumour cells might thus aid the development of new anticancer treatments. Writing in *Cell*, Hsu *et al.*[4] shed light on a previously unknown aspect of how chemotherapy-induced entry into senescence is controlled.

Although much progress has been made in uncovering factors that drive senescence, the processes that ultimately commit cells to this fate are poorly understood. A growing body

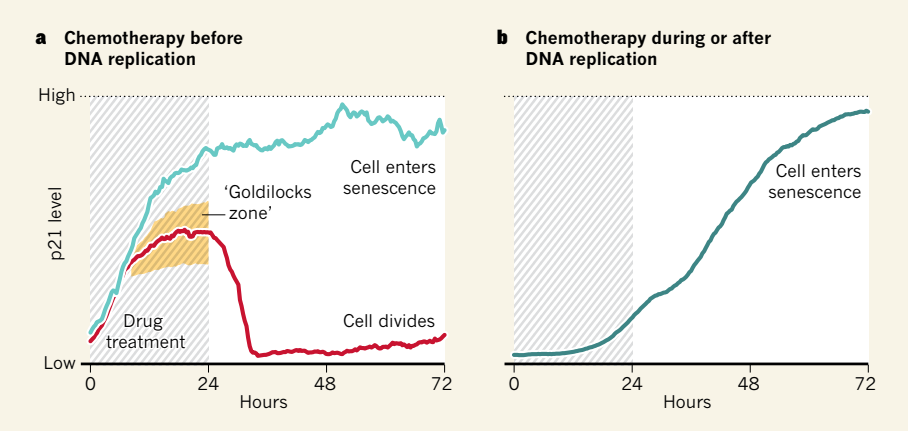


Figure 1 | Levels of p21 protein and cancer-cell fate after chemotherapy. **a**, Hsu *et al.*[4] found that, when human cancer cells grown *in vitro* were treated with chemotherapy drugs at a stage in the cell cycle before DNA replication, two types of cell fate were observed. In some cells, p21 levels rose and the cells entered a permanent state of non-division termed senescence. In other cells, after an initial rise in the level of p21, the protein returned to a low level and the cells divided. The authors describe this p21-dependent switch in cell fate as being affected by a 'Goldilocks zone' (yellow), in which the levels and dynamics of the protein after drug treatment must be 'just right' for cells to be able to halt the cell cycle, repair DNA and then continue to divide. **b**, By contrast, if cells received drug treatment during or after DNA replication, p21 levels gradually rose and cells entered senescence. (Graphs based on Fig. 3 of ref. 4, showing just the 72-hour window during and after drug treatment.)

of evidence indicates that the regulation of commitment to enter senescence is complex. The mere presence of factors associated with triggering this cell fate is not in itself sufficient to provide an 'on switch' for senescence.

The protein p21 is probably best known for its role in blocking cell division by inhibiting protein complexes called cyclin-dependent kinases. If DNA damage occurs, p21 activity halts cell division and growth[5], giving cells time for DNA repair and thereby preventing such damage from having catastrophic cellular consequences. There is evidence that p21 can induce senescence during chemotherapy[6]. Yet, paradoxically, some research suggests that the protein can promote cancer-cell division after chemotherapy[7]. One possible explanation for this discrepancy is that the abundance and dynamics of p21 after chemotherapy have a key role in determining whether cancer cells enter senescence or divide.

To test this idea, Hsu and colleagues developed a microscopy system to study thousands of individual, cultured human lung and colon cancer cells that had been treated with a DNA-damaging chemotherapy drug. The authors monitored the abundance of p21 by tagging it with a fluorescent protein, and also tracked the progression of the various stages of the cell cycle. In contrast to previous research suggesting that high levels of p21 invariably lead to either cell growth or senescence[5,7], the authors describe a complex, but unifying, picture of how p21 levels relate to cell fate. Hsu *et al.* noted that if chemotherapy resulted in an initial rise in p21 levels followed by a decline to low levels, cell division, rather than senescence, occurred (Fig. 1). Cancer cells that entered senescence after drug treatment initially had a low level of p21 that gradually rose to a high level.

Hsu and colleagues suggest that there is a 'Goldilocks zone' for proliferation — a level of p21 that is 'just right' to allow tumour cells to divide after chemotherapy. How might p21 dynamics control cell fate in this way? Chemotherapy drugs are most damaging to DNA if given to cells at the cell-cycle stage at which DNA replication occurs[8]. It might therefore be expected that cells given chemotherapy during DNA replication would have higher levels of p21 than would cells treated before DNA replication occurs. Yet, surprisingly, Hsu and colleagues found that cells treated during DNA replication had high levels of DNA damage but low levels of p21, and that levels of p21 then increased over time. By contrast, drug treatment before DNA replication resulted in a rapid rise in p21 expression that, depending on the individual cell, either returned to a low level or rose further.

How do some cancer cells that have undergone drug-induced DNA damage revert to having low levels of p21 expression and gain the ability to divide? The authors propose a model that incorporates dynamic regulation of p21 expression and the level of DNA damage. They suggest that cell fate after chemotherapy shows a property termed bistability — cells are poised to follow one of two fates.

In this scenario, at the cell-cycle stage before DNA replication, if cells express intermediate levels of p21 and small fluctuations occur in signals identifying DNA damage due to chemotherapy, such fluctuations might promote either the rapid induction or decline of p21 expression, driving cells to, respectively, enter senescence or divide.

However, when cells undergo DNA replication, a stage of the cell cycle at which DNA-damage signals are higher than normal (because errors can occur during DNA replication), only a slight increase in the level of p21 would be enough to establish a stable state of high p21 that would lead to senescence. Thus, the Goldilocks zone is defined by the level of p21 and DNA damage, and determines whether cells divide after chemotherapy. To assess the clinical relevance of this finding, the dose and drug dependency of p21 dynamics during chemotherapy should be examined in detail.

These results raise the possibility that targeting cancer cells at specific cell-cycle stages would produce major differences in the cellular response to chemotherapy, with cells targeted during or shortly after DNA replication being more likely to enter senescence than cells targeted before replication. This model should be investigated further. Taking such an approach in the clinic would pose challenges, however, given that tumours contain a mixture of cell populations that are at different cell-cycle stages. One strategy to tackle this could be to direct cells towards the specific cell-cycle stage at which chemotherapy would be most effective. The authors found that if cells were treated with a small molecule that triggers DNA replication, senescence occurred more commonly than cell division after chemotherapy.

Another challenge will be to identify the optimal outcome for a given cancer following chemotherapy. Although senescence might be an ideal response for certain tumours, others that are more prone to dying in response to cellular damage might be more effectively treated by inducing cell death rather than by triggering senescence.

Hsu and colleagues' work provides a detailed foundation for understanding what governs the fate of cancer cells after chemotherapy. Now it is time to build on this progress, to determine whether strategies can be found that maximize the effectiveness of our current arsenal of anticancer agents. ■

Yunpeng Liu *and* **Michael T. Hemann**
are at the Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.
e-mails: hemann@mit.edu; yunpengl@mit.edu

1. Campisi, J. & d'Adda di Fagagna, F. *Nature Rev. Mol. Cell Biol.* **8**, 729–740 (2007).
2. Campisi, J. *Annu. Rev. Physiol.* **75**, 685–705 (2013).
3. Ewald, J. A., Desotelle, J. A., Wilding, G. & Jarrard, D. F. *J. Natl Cancer Inst.* **102**, 1536–1546 (2010).
4. Hsu, C.-H., Altschuler, S. J. & Wu, L. F. *Cell* **178**, 361–373 (2019).
5. d'Adda di Fagagna, F. *Nature Rev. Cancer* **8**, 512–522 (2008).
6. Fitzgerald, A. L. *et al. Cell Death Dis.* **6**, e1678 (2015).
7. Abbas, T. & Dutta, A. *Nature Rev. Cancer* **9**, 400–414 (2009).
8. Potter, A. J. *et al. Carcinogenesis* **23**, 389–401 (2002).

This article was published online on 5 August 2019.

ARTICLE

https://doi.org/10.1038/s41586-019-1457-z

Exome sequencing of Finnish isolates enhances rare-variant association power

Adam E. Locke[1,2,3,43], Karyn Meltz Steinberg[2,4,43], Charleston W. K. Chiang[5,6,7,43], Susan K. Service[5,43], Aki S. Havulinna[8,9], Laurel Stell[10], Matti Pirinen[8,11,12], Haley J. Abel[2,13], Colby C. Chiang[2], Robert S. Fulton[2], Anne U. Jackson[3], Chul Joo Kang[2], Krishna L. Kanchi[2], Daniel C. Koboldt[2,14,15], David E. Larson[2,13], Joanne Nelson[2], Thomas J. Nicholas[2,16], Arto Pietilä[9], Vasily Ramensky[5,17], Debashree Ray[3,18], Laura J. Scott[3], Heather M. Stringham[3], Jagadish Vangipurapu[19], Ryan Welch[3], Pranav Yajnik[3], Xianyong Yin[3], Johan G. Eriksson[20,21,22], Mika Ala-Korpela[23,24,25,26,27,28], Marjo-Riitta Järvelin[29,30,31,32,33], Minna Männikkö[30,34], Hannele Laivuori[7,35,36], FinnGen Project[37]Susan K. Dutcher[2,13], Nathan O. Stitziel[2,38], Richard K. Wilson[2,14,15], Ira M. Hall[1,2], Chiara Sabatti[9,39], Aarno Palotie[7,40,41], Veikko Salomaa[9], Markku Laakso[19,42], Samuli Ripatti[7,11,41], Michael Boehnke[3,44]* & Nelson B. Freimer[5,44]*

Exome-sequencing studies have generally been underpowered to identify deleterious alleles with a large effect on complex traits as such alleles are mostly rare. Because the population of northern and eastern Finland has expanded considerably and in isolation following a series of bottlenecks, individuals of these populations have numerous deleterious alleles at a relatively high frequency. Here, using exome sequencing of nearly 20,000 individuals from these regions, we investigate the role of rare coding variants in clinically relevant quantitative cardiometabolic traits. Exome-wide association studies for 64 quantitative traits identified 26 newly associated deleterious alleles. Of these 26 alleles, 19 are either unique to or more than 20 times more frequent in Finnish individuals than in other Europeans and show geographical clustering comparable to Mendelian disease mutations that are characteristic of the Finnish population. We estimate that sequencing studies of populations without this unique history would require hundreds of thousands to millions of participants to achieve comparable association power.

Most alleles with demonstrated deleterious effects on phenotypes directly alter the structure or function of a protein[1,2]. Exome-sequencing studies aim to discover such alleles and demonstrate their association to common diseases and disease-related quantitative traits. However, exome-sequencing studies to date generally have identified few newly associated rare variants or genes[3,4]. The sample size that is required for such discoveries remains uncertain and theoretical analyses indicate that studies to date have been underpowered, as most deleterious variants are expected to be rare owing to purifying selection[5]. These previous analyses also suggest that the power to detect associations to deleterious alleles is highest in populations that have expanded in isolation after recent bottlenecks, as alleles passing through the bottlenecks may increase to much higher frequencies than in other populations[6–8].

Finland exemplifies such a history. Bottlenecks occurred at the founding of early-settlement regions (southern and western Finland) 2,000–4,000 years ago and again with internal migration to late-settlement regions (northern and eastern Finland) in the fifteenth and sixteenth centuries[9]. Finland's subsequent population growth (to approximately 5.5 million) generated sizable geographical sub-isolates in late-settlement regions.

This unique population history has resulted in 'the Finnish Disease Heritage'[10], 36 Mendelian diseases that are much more common in

[1]Department of Medicine, Washington University School of Medicine, St Louis, MO, USA. [2]McDonnell Genome Institute, Washington University School of Medicine, St Louis, MO, USA. [3]Department of Biostatistics and Center for Statistical Genetics, University of Michigan School of Public Health, Ann Arbor, MI, USA. [4]Department of Pediatrics, Washington University School of Medicine, St Louis, MO, USA. [5]Center for Neurobehavioral Genetics, Jane and Terry Semel Institute for Neuroscience and Human Behavior, University of California Los Angeles, Los Angeles, CA, USA. [6]Center for Genetic Epidemiology, Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA. [7]Quantitative and Computational Biology Section, Department of Biological Sciences, University of Southern California, Los Angeles, CA, USA. [8]Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland. [9]National Institute for Health and Welfare, Helsinki, Finland. [10]Department of Biomedical Data Science, Stanford University, Stanford, CA, USA. [11]Department of Public Health, University of Helsinki, Helsinki, Finland. [12]Helsinki Institute for Information Technology HIIT and Department of Mathematics and Statistics, University of Helsinki, Helsinki, Finland. [13]Department of Genetics, Washington University School of Medicine, St Louis, MO, USA. [14]The Institute for Genomic Medicine, Nationwide Children's Hospital, Columbus, OH, USA. [15]Department of Pediatrics, The Ohio State University College of Medicine, Columbus, OH, USA. [16]USTAR Center for Genetic Discovery and Department of Human Genetics, University of Utah, Salt Lake City, UT, USA. [17]Federal State Institution "National Medical Research Center for Preventive Medicine" of the Ministry of Healthcare of the Russian Federation, Moscow, Russia. [18]Departments of Epidemiology and Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA. [19]Institute of Clinical Medicine, Internal Medicine, University of Eastern Finland, Kuopio, Finland. [20]Department of Public Health Solutions, National Institute for Health and Welfare, Helsinki, Finland. [21]Folkhälsan Research Center, Helsinki, Finland. [22]Department of General Practice and Primary Health Care, University of Helsinki, Helsinki and Helsinki University Hospital, Helsinki, Finland. [23]Systems Epidemiology, Baker Heart and Diabetes Institute, Melbourne, Victoria, Australia. [24]Computational Medicine, Faculty of Medicine, University of Oulu and Biocenter Oulu, University of Oulu, Oulu, Finland. [25]NMR Metabolomics Laboratory, School of Pharmacy, University of Eastern Finland, Kuopio, Finland. [26]Population Health Science, Bristol Medical School, University of Bristol, Bristol, UK. [27]Medical Research Council Integrative Epidemiology Unit at the University of Bristol, Bristol, UK. [28]Department of Epidemiology and Preventive Medicine, School of Public Health and Preventive Medicine, Faculty of Medicine, Nursing and Health Sciences, The Alfred Hospital, Monash University, Melbourne, Victoria, Australia. [29]Biocenter Oulu, University of Oulu, Oulu, Finland. [30]Center for Life Course Health Research, Faculty of Medicine, University of Oulu, Oulu, Finland. [31]Unit of Primary Health Care, Oulu University Hospital, Oulu, Finland. [32]Department of Epidemiology and Biostatistics, MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, London, UK. [33]Department of Life Sciences, College of Health and Life Sciences, Brunel University London, London, UK. [34]Northern Finland Birth Cohorts, Faculty of Medicine, University of Oulu, Oulu, Finland. [35]Medical and Clinical Genetics, University of Helsinki and Helsinki University Hospital, Helsinki, Finland. [36]Department of Obstetrics and Gynecology, Tampere University Hospital and University of Tampere, Faculty of Medicine and Life Sciences, Tampere, Finland. [37]A list of participants and their affiliations appears in the Supplementary Information. [38]Cardiovascular Division, Department of Medicine, Washington University School of Medicine, St Louis, MO, USA. [39]Department of Statistics, Stanford University, Stanford, CA, USA. [40]Analytical and Translational Genetics Unit (ATGU), Psychiatric & Neurodevelopmental Genetics Unit, Departments of Psychiatry and Neurology, Massachusetts General Hospital, Boston, MA, USA. [41]Broad Institute of MIT and Harvard, Cambridge, MA, USA. [42]Department of Medicine, Kuopio University Hospital, Kuopio, Finland. [43]These authors contributed equally: Adam E. Locke, Karyn Meltz Steinberg, Charleston W. K. Chiang, Susan K. Service. [44]These authors jointly supervised this work: Michael Boehnke, Nelson B. Freimer. *e-mail: boehnke@umich.edu; nfreimer@mednet.ucla.edu

Finnish individuals than in other Europeans. These disorders concentrate in late-settlement regions of Finland[10], and the genes responsible for them exhibit extreme enrichment of deleterious variants[11–13]. We created the Finnish Metabolic Sequencing (FinMetSeq) study to capitalize on the population history of late-settlement Finland to discover rare-variant associations with cardiovascular and metabolic disease-relevant quantitative traits through exome sequencing of two extensively phenotyped population cohorts, FINRISK and METSIM (Methods).

We successfully sequenced 19,292 FinMetSeq participants and tested the identified variants for association with 64 clinically relevant quantitative traits, discovering 43 novel associations with deleterious variants[14,15]: 19 associations (11 traits) in FinMetSeq alone and 24 associations (20 traits) in a combined analysis of FinMetSeq with 24,776 Finns from three cohorts with imputed genome-wide genotypes. Of the 26 variants that underlie these 43 associations, 19 were unique to Finland or enriched more than 20-fold in FinMetSeq compared to non-Finnish Europeans (NFE). These enriched alleles cluster geographically like Finnish Disease Heritage mutations, indicating that the distribution of trait-associated rare alleles may vary significantly between locations within a country.

We demonstrate that exome sequencing in a historically isolated population that expanded after recent population bottlenecks is an efficient strategy to discover alleles with a substantial effect on quantitative traits. As most of the novel, putatively deleterious trait-associated variants that we identified are unique to or highly enriched in Finland, we estimate that similarly powered studies of these variants in non-Finnish populations would require hundreds of thousands or millions of participants.

Genetic variation

In 19,292 successfully sequenced exomes, we identified 1,318,781 single-nucleotide variants and 92,776 insertion or deletion variants (Supplementary Tables 1–3 and Supplementary Information). Compared to NFE control exomes (gnomAD v.2.1, Extended Data Fig. 1a), FinMetSeq exomes showed depletion of singletons and doubletons and excess variants with minor allele count (MAC) ≥ 5, particularly for predicted-deleterious alleles (Extended Data Fig. 1b).

Association analyses

We tested for association between genetic variants in FinMetSeq and 64 clinically relevant quantitative traits after standard adjustments for medications and covariates, and transformation to normality for analyses (Methods, Supplementary Tables 4, 5). Out of 64 traits, 62 exhibited significant heritability with common single-nucleotide variants ($P < 0.05$; $5\% < h^2 < 53\%$; Extended Data Fig. 2a, Supplementary Table 6), with substantial phenotypic and genetic correlations between traits (Extended Data Fig. 2b).

Single-variant association tests with genetic variants with MAC ≥ 3 among the 3,558 to 19,291 individuals measured for each trait (Supplementary Tables 4, 5) identified 1,249 associations ($P < 5 \times 10^{-7}$) at 531 variants (Supplementary Table 7); 53 traits were associated with at least one variant (Fig. 1a). All 1,249 associations remained significant after adjustment for multiple testing (exome-wide and across the 64 traits using a hierarchical procedure setting average the false discovery rate (FDR) to 5%; see Methods). Using this procedure on the 531 associated variants, we detected 287 more associations (Supplementary Table 8), most of which reflected a high correlation between lipid traits. Of the 531 variants, those with a greater than 10× frequency in FinMetSeq compared to NFE were more likely to be trait-associated (odds ratio = 4.92, $P = 2.6 \times 10^{-5}$; Extended Data Fig. 1c).

After clumping associated variants within 1 megabase (Mb) and with $r^2 > 0.5$ into single loci (Methods), the 531 associated variants represented 262 distinct loci (597 trait–locus pairs; Supplementary Table 7). The number of associated loci per trait correlated positively with trait heritability ($r = 0.38$, $P = 8.8 \times 10^{-4}$), although height was a notable outlier (Fig. 1b).

Most variants and loci (61%) were associated with a single trait; 4% were associated with ≥10 traits. Overlapping associations (Extended Data Fig. 3a) reflect both phenotypic and genetic correlations and the estimated genetic correlation of trait pairs predicts shared loci between traits (Extended Data Fig. 3b). Gene-based association tests revealed 54 associations with $P < 3.88 \times 10^{-6}$ and multi-trait FDR-corrected $P < 0.05$ (Methods and Supplementary Table 9), including 10 traits associated with *APOB* (Extended Data Fig. 4) and a novel association of *SECTM1* with high density lipoprotein cholesterol subfraction 2 (HDL2-C) (Extended Data Fig. 5).

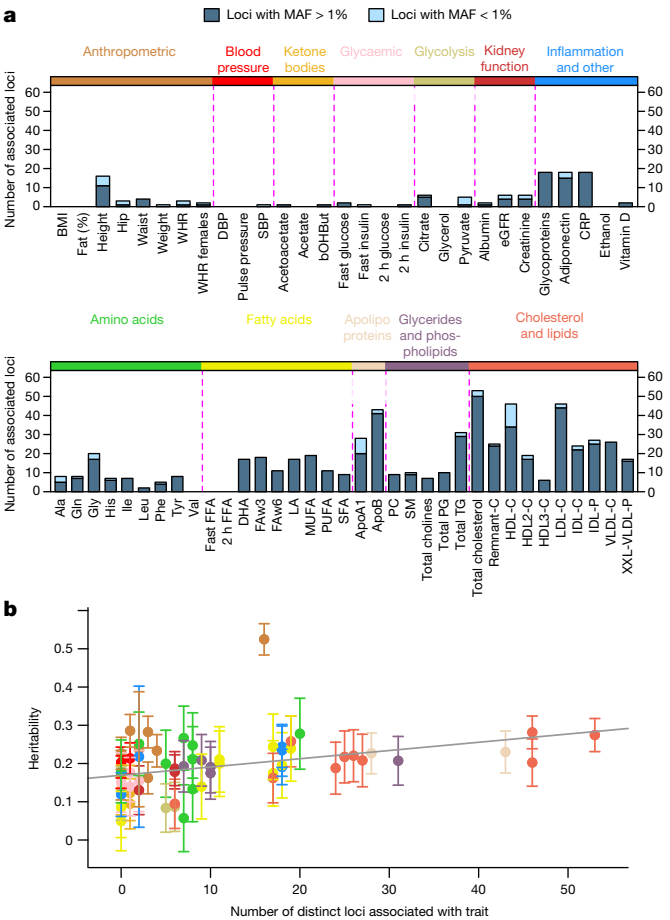


Fig. 1 | Characterization of associations. a, Numbers of genomic loci associated with each trait. Bars are subdivided into common (MAF > 1%, dark blue) and rare (MAF ≤ 1%, light blue) variants. **b**, Relationship between estimated heritability and number of loci detected per trait. Each trait is coloured by trait group. Data are mean ± s.e.m. The grey line shows the linear regression fit to indicate the general trend. The number of independent individuals used in each point is listed in Supplementary Table 5. Height is the notable outlier. See Supplementary Table 4 for abbreviations.

To determine which of the 1,249 single-variant associations are distinct from previous GWAS findings, we repeated the association analysis for each trait conditioning on published associated variants in the EBI GWAS Catalog (as per December 2016, Methods); 478 associations at 126 loci remained significant ($P < 5 \times 10^{-7}$), including at least one association for 48 traits (Supplementary Table 10). Conditionally associated variants were more often rare (24% versus 11%), more likely protein-altering (31% versus 22%) and more frequently >10× enriched in FinMetSeq relative to NFE (19% versus 10%) than associated variants overall.

Replication and follow-up

We attempted to replicate the 478 single-variant associations (unconditional and conditional $P \leq 5 \times 10^{-7}$) and follow up on 2,120 sub-threshold associations from FinMetSeq (unconditional $5 \times 10^{-7} < P \leq 5 \times 10^{-5}$ and conditional $P \leq 5 \times 10^{-5}$) in 24,776

Table 1 | Associations with predicted deleterious variants from FinMetSeq or combined analysis

Chromosome: position	Gene	FinMetSeq MAF	NFE MAF	MAF ratio (95% CI)	Trait	FinMetSeq P	FinMetSeq β	Replication or combined *P*	Replication or combined β
1:55,076,137	*FAM151A*	0.099	0.0147	6.7 (6.1–7.5)	IDL-C	5.4×10^{-16}	−0.187	**2.1×10^{-17}**	−0.191
					IDL-P	8.9×10^{-14}	−0.172	**1.9×10^{-16}**	−0.185
2:120,848,049	*EPB41L5*	0.085	0.044	1.9 (1.8–2.1)	eGFR[a]	1.7×10^{-6}	−0.093	4.8×10^{-12}	−0.107
					Creatinine[a]	2.5×10^{-6}	0.091	2.5×10^{-12}	0.098
3:125,831,672	*ALDH1L1*	0.0026	0	∞	Gly	1.8×10^{-8}	−0.873	**4.5×10^{-4}**	−0.827
4:13,612,630	*BOD1L1*	0.0001	0	∞	WHR	4.7×10^{-7}	−2.501	NA	NA
5:79,336,091	*THBS4*	0.0045	0.0001	45 (14.4–140.9)	Weight[a]	6.7×10^{-7}	−0.377	3.2×10^{-7}	−0.252
5:140,181,423	*PCDHA3*	0.0001	NA	NA	WHR	2.7×10^{-7}	2.559	NA	NA
9:107,548,661	*ABCA1*	0.00023	0	∞	HDL-C	4.8×10^{-10}	−2.046	NA	NA
9:136,501,728	*DBH*	0.05	0.0021	23.8 (18.4–30.4)	DBP[a]	1.5×10^{-6}	−0.115	2.8×10^{-12}	−0.11
11:47,282,929	*NR1H3*	0.0042	0.00003	140 (19.5–1004.4)	HDL-C	1.4×10^{-7}	0.425	**6.7×10^{-7}**	0.435
					HDL2-C[a]	3.2×10^{-6}	0.473	1.3×10^{-8}	0.458
					VLDL-C[a]	4.0×10^{-6}	−0.469	3.1×10^{-7}	−0.412
11:116,692,293	*APOA4*	0.0096	0.012	0.8 (0.7–0.9)	HDL-C[a]	2.2×10^{-5}	0.225	1.5×10^{-7}	0.196
11:117,352,857	*DSCAML1*	0.016	0.0002	80 (35.7–179.3)	VLDL-C	4.1×10^{-8}	0.299	**2.0×10^{-3}**	0.162
14:101,198,426	*DLK1*	0.023	0.00013	177 (66.3–472.4)	Height[a]	2.7×10^{-5}	−0.149	1.2×10^{-10}	−0.163
16:55,862,682	*CES1*	0.0018	0.00003	60 (8.3–432.0)	HDL-C	1.1×10^{-10}	0.771	**3.8×10^{-6}**	0.793
					ApoA1[a]	1.9×10^{-6}	0.668	4.0×10^{-9}	0.718
16:56,996,009	*CETP*	0.0017	0.00003	56.7 (7.9–408.3)	ApoA1	2.6×10^{-8}	0.834	**1.8×10^{-4}**	1.034
					HDL-C	1.1×10^{-14}	0.946	**8.8×10^{-21}**	1.217
16:68,013,570	*DPEP3*	0.0099	0.00044	22.5 (12.9–39.1)	HDL-C	1.6×10^{-7}	−0.295	**7.2×10^{-15}**	−0.373
					ApoA1[a]	5.2×10^{-6}	−0.294	4.0×10^{-7}	−0.253
16:68,732,169	*CDH3*	0.0044	0.00064	6.9 (4.2–11.2)	Pyruvate[a]	3.7×10^{-5}	0.417	6.6×10^{-10}	0.471
17:6,599,157	*SLC13A5*	0.00091	0	∞	Citrate	1.3×10^{-9}	1.294	**9.5×10^{-12}**	1.309
17:7,129,898	*DVL2*	0.02	0.02	1.0 (0.9–1.1)	Val[a]	4.2×10^{-5}	−0.239	5.7×10^{-9}	−0.232
17:39,135,270	*KRT40*	0.00013	0	∞	HDL-C	3.2×10^{-8}	2.416	NA	NA
17:41,062,979	*G6PC*	0.025	0	∞	MUFA	4.4×10^{-7}	0.275	3.5×10^{-1}	0.067
					Glycerol[a]	5.8×10^{-6}	0.218	4.1×10^{-7}	0.183
					CRP[a]	1.6×10^{-5}	0.175	4.0×10^{-9}	0.185
					Total TG[a]	1.0×10^{-6}	0.23	1.3×10^{-7}	0.197
17:41,926,216	*CD300LG*	0.00034	0	∞	HDL-C	4.8×10^{-14}	2.061	**4.9×10^{-2}**	0.801
					HDL2-C	1.3×10^{-7}	2.154	NA	NA
					ApoA1	8.1×10^{-8}	1.694	NA	NA
18:47,091,686	*LIPG*	0.0025	0	∞	HDL2-C[a]	1.2×10^{-5}	0.579	5.6×10^{-10}	0.624
					PC[a]	3.1×10^{-6}	0.624	1.1×10^{-8}	0.578
					Total PG[a]	9.0×10^{-6}	0.594	1.1×10^{-7}	0.538
19:10,683,762	*AP1M2*	0.015	0.00009	167 (41.6–668.5)	ApoB	5.8×10^{-8}	−0.282	**1.5×10^{-3}**	−0.199
					IDL-C[a]	1.1×10^{-6}	−0.289	6.9×10^{-14}	−0.319
					IDL-P[a]	2.1×10^{-6}	−0.281	8.5×10^{-14}	−0.318
					Remnant-C[a]	8.0×10^{-6}	−0.268	2.7×10^{-12}	−0.301
19:11,350,904	*ANGPTL8*	0.0025	0	∞	HDL2-C[a]	3.4×10^{-6}	0.564	1.1×10^{-8}	0.574
19:49,318,380	*HSD17B14*	0.046	0.05	0.9 (0.8–1.0)	Val[a]	3.4×10^{-5}	−0.152	2.1×10^{-7}	−0.144
20:24,994,201	*ACSS1*	0.0026	0	∞	Acetate[a]	1.3×10^{-5}	0.626	2.1×10^{-12}	0.631

Chromosome positions were based on GRCh37. NFE MAFs were taken from gnomAD v.2.1 control exomes excluding Estonian or Swedish individuals. MAF: 0, variant present in gnomAD, but not in NFE controls; NA, variant not present in gnomAD. Replication values with $P < 0.05$ are highlighted in bold. 95% CI, 95% confidence interval. See Supplementary Table 4 for trait abbreviations.
[a]Associated traits that only reach significance in combined analysis.

participants from three Finnish cohort studies: FINRISK[16,17] participants not in FinMetSeq ($n = 18{,}215$), Northern Finland Birth Cohort 1966[18] ($n = 5{,}139$) and Helsinki Birth Cohort[19] ($n = 1{,}412$), all imputed using the Finnish SISu v.2 reference panel (www.sisuproject.fi). Following association analysis within each cohort, we conducted a meta-analysis of the three imputation-based studies to test for replication of FinMetSeq variants (replication analysis) and a four-study meta-analysis with FinMetSeq to follow up on suggestive associations (combined analysis).

Of 448 significant variant–trait associations with replication data, 392 (87.5%) replicated at $P < 0.05$ (Supplementary Table 11). Of the 1,417 sub-threshold associations, 431 reached $P < 5 \times 10^{-7}$ in the combined analysis (Supplementary Table 12); more than 60% of the variants were absent from the reference panel and thus could not be tested further.

Among the significant associations from FinMetSeq or the combined analysis, 43 associations were with 26 predicted deleterious variants (6 protein truncating variants (PTVs) and 20 missense variants) that

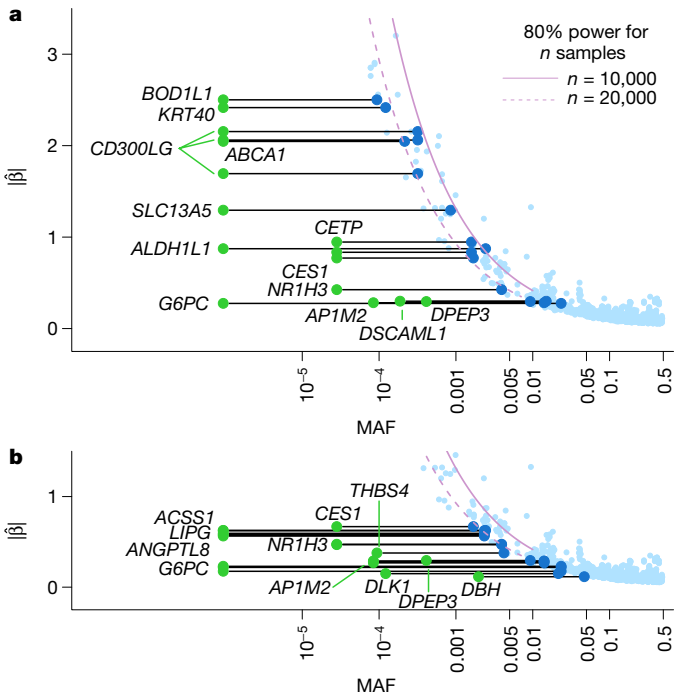


Fig. 2 | Allelic enrichment in the Finnish population and its effect on genetic discovery. a, Relationship between MAF and estimated effect size for associations discovered in FinMetSeq. Each variant that reached significance in FinMetSeq was plotted, with associations in Table 1 represented by dark-blue points (FinMetSeq MAFs) and green points (NFE MAFs). Purple lines indicate 80% power curves for sample sizes of $n = 10{,}000$ and $n = 20{,}000$ at $\alpha = 5 \times 10^{-7}$. **b**, Same plot as in **a**, highlighting the variants in Table 1 that only reached significance in the combined analysis.

conditional analysis and literature review suggest are novel (Table 1). Of those, 19 associations (15 variants) were significant in FinMetSeq (Table 1 and Supplementary Table 11); another 24 associations (16 variants) reached significance in the combined analysis (Table 1 and Supplementary Table 12). Furthermore, 34 out of 43 associations were with 19 variants either found only in Finland or enriched more than 20-fold in FinMetSeq compared to NFE. The identification of associations for these 19 variants would have required much larger samples in NFE populations than in FinMetSeq (Fig. 2a, b). We provide brief summaries relating some of these associations to known biology and previously described genetic evidence (Table 1, expanded version in Supplementary Table 13; see Supplementary Information), highlighting here the most notable findings.

Anthropometric traits

A predicted damaging missense variant (Arg94Cys) in *THBS4*, which was 45× more frequent in FinMetSeq than in NFE, was associated in the combined analysis with a mean 5.9 kg decrease in body weight. *THBS4* encodes thrombospondin 4, a matricellular protein that is found in blood vessel walls and highly expressed in heart and adipose tissues[20]. THBS4 may regulate vascular inflammation[21] and has been implicated in the risk of heart disease[22].

A predicted damaging missense variant (Val104Met) in *DLK1*, which was 177× more frequent in FinMetSeq than in NFE, was associated in the combined analysis with a mean 1.3 cm decrease in height. *DLK1* encodes delta-like notch ligand 1, an epidermal growth factor that interacts with fibronectin and inhibits adipocyte differentiation. Uniparental disomy of *DLK1* causes Temple and Kagami–Ogata syndromes, which are characterized by growth restriction, hypotonia, joint laxity, motor delay and early onset of puberty[23]. Paternally inherited common variants near *DLK1* are associated with childhood obesity, type 1 diabetes, age at menarche and precocious puberty[24–26]. Homozygous null mutations in the mouse orthologue *Dlk1* lead to embryos with reduced size, skeletal length and lean mass[27]; in Darwin's finches, single-nucleotide variants at this locus have a strong effect on beak size[28].

High-density lipoprotein cholesterol

A predicted deleterious missense variant (Arg112Trp) in *CD300LG* is associated in FinMetSeq with a mean 0.95 mmol l^{-1} increase in high-density lipoprotein cholesterol (HDL-C) and is associated with increased HDL2-C and ApoA1. This variant, which is absent from NFE, has an opposite direction of effect from a previously reported deleterious missense variant in this gene[29], which encodes a type-I cell-surface glycoprotein.

Amino acids

A stop gain variant (Arg722X) in *ALDH1L1* is associated in FinMetSeq with reduced serum glycine levels and is absent from NFE; this trait may increase risk for cardiometabolic disorders[30,31]. *ALDH1L1* encodes 10-formyltetrahydrofolate dehydrogenase, which competes with serine hydroxymethyltransferase to alter the ratio of serine to glycine in the cytosol. Gene-based tests suggest that additional PTVs and missense variants in *ALDH1L1* alter glycine levels ($P = 1.4 \times 10^{-20}$; Extended Data Fig. 6 and Supplementary Table 9).

Ketone bodies

A predicted damaging missense variant (Phe517Ser) in *ACSS1* is associated in the combined analysis with increased serum acetate levels and is absent from NFE. *ACSS1* encodes an acyl-coenzyme A synthetase and has a role in the conversion of acetate to acetyl-CoA. In rodents, increased acetate levels lead to obesity, insulin resistance and metabolic syndrome[32].

Trait–associations and disease end points

Genotype data from FinnGen[33] enabled us to test whether deleterious variants responsible for our novel trait associations contributed to related disease end points. We examined 22 diseases for the 25 available variants shown in Table 1; 3 variants were associated with diseases in FinnGen at a Bonferroni threshold value of $P < 0.05/(22 \times 25) = 9.0 \times 10^{-5}$ (Supplementary Table 14).

A predicted damaging missense variant (Ser32Pro) in *KRT40*, which is associated in FinMetSeq with elevated HDL-C but is absent in NFE, is associated in FinnGen with increased risk of pancreatitis. Although this is the first disease association reported for *KRT40*, type-I keratins regulate exocrine pancreas homeostasis[34]. A 29-bp deletion that causes a frameshift in *FAM151A* is associated in FinMetSeq with decreased total cholesterol in intermediate-density lipoproteins (IDL-C) and decreased concentration of IDL particles, is 6.7× more frequent in FinMetSeq than NFE and is associated in FinnGen with decreased risk of myocardial infarction. Interpretation of this association is complicated as the variant is also situated in an overlapping gene (*ACOT11*), which is involved in fatty acid metabolism and lies <1Mb from a cardioprotective variant in *PCSK9*. Finally, a predicted damaging missense variant (Arg65Trp) in *DBH*, which is associated with a mean 1.0 mm Hg decrease in diastolic blood pressure in the combined analysis, is 23.8× more frequent in FinMetSeq than in NFE, and is associated in FinnGen with decreased risk of hypertension. Distinct loci in this gene and gene-based tests are associated with mean arterial pressure[35,36].

Replication outside Finland

To assess the generalizability of these novel associations, we attempted to replicate associations from our combined analysis with data from the UK Biobank. Across 8 anthropometric and blood pressure traits for which UK Biobank data are publicly available, our combined analysis identified 31 trait–variant associations, of which 23 were present in the UK Biobank. Of the 23 associations, 20 were to variants with a minor allele frequency (MAF) > 1% in FinMetSeq and a comparable frequency in UK Biobank; 15 (75%) showed association in UK Biobank at $P < 0.05/23 = 2.2 \times 10^{-3}$. The three rare variants in this

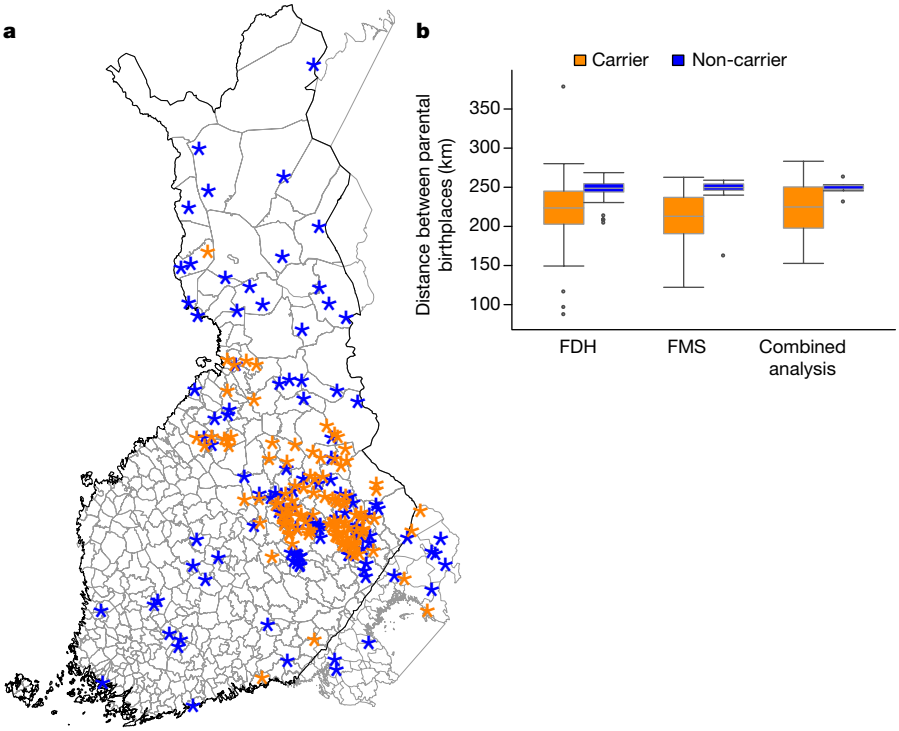


Fig. 3 | Geographical clustering of associated variants. a, Example of geographical clustering for a novel trait-associated variant (Table 1). The map shows birth locations of all 113 parents of carriers (orange) and 113 randomly selected parents of non-carriers (blue) of the minor allele for rs780671030 in *ALDH1L1*. **b**, Mutations in the Finnish Disease Heritage (FDH) genes ($n = 38$) geographically cluster (by parental birthplace) similarly to trait-associated variants (Table 1) that are >10× more frequent in FinMetSeq than in NFE ($n = 12$) and more than enriched variants from our combined analysis ($n = 7$). For all variants, carriers clustered more than non-carriers (centre line, median; box limits, upper and lower quartiles; whiskers, 1.5× interquartile range; points, outliers). Birthplaces of carrier and non-carrier individuals were plotted on a map of Finland, including regions that were ceded before the Second World War (© Karttakeskus Oy, 2001).

analysis were all more than 10× more frequent in FinMetSeq than in UK Biobank; none were associated in UK Biobank (Supplementary Table 15). However, even after adjusting for winner's curse[37], we had <50% power to detect these associations in UK Biobank, consistent with the argument that extremely large samples will be needed in other populations to achieve the power for rare-variant association studies that we observed in Finland.

Enriched variants cluster geographically

Given the concentration of Finnish Disease Heritage mutations within regions of late-settlement Finland[38], we hypothesized that trait-associated variants discovered through FinMetSeq would also cluster geographically. Principal component analysis supported this hypothesis, revealing a broad-scale population structure within late-settlement regions among 14,874 unrelated FinMetSeq participants with known parental birthplaces (Extended Data Fig. 7). Carriers of PTVs and missense alleles showed more clustering of parental birthplaces than carriers of synonymous alleles, even after adjusting for MAC (Supplementary Table 16a, b).

To analyse the distribution of variants within late-settlement Finland, we delineated geographically distinct population clusters using haplotype sharing among 2,644 unrelated individuals with both parents born in the same municipality (Methods and Extended Data Fig. 8). We compared variant counts across functional classes and frequencies between an early-settlement reference cluster and 12 clusters containing ≥100 individuals (Extended Data Fig. 9 and Supplementary Tables 17, 18). Clusters that represent the most heavily bottlenecked late-settlement regions (Lapland and Northern Ostrobothnia) displayed a deficit of singletons and enrichment of intermediate frequency variants compared to other clusters.

Variants that were more than 10× enriched in FinMetSeq compared to NFE displayed particularly strong geographical clustering (Supplementary Table 19). We further characterized clustering for FinMetSeq-enriched trait-associated variants, by comparing mean distances between birthplaces of parents of minor allele carriers to those of non-carriers (Supplementary Table 20). Most of these variants were highly localized. For example, for rs780671030 in *ALDH1L1*, the mean distance between parental birthplaces is 135 km for carriers and 250 km for non-carriers ($P < 1.0 \times 10^{-7}$, Fig. 3a).

Finally, we identified comparable geographical clustering between carriers of 35 Finnish Disease Heritage mutations and carriers of FinMetSeq-enriched trait-associated variants (Fig. 3b and Methods). Clustering was considerably greater in carriers than clustering observed for non-carriers of both sets of variants, suggesting that rare trait-associated variants may be much more unevenly distributed geographically than has previously been appreciated.

Discussion

We demonstrate that a well-powered exome-sequencing study of deeply phenotyped individuals can identify numerous rare variants that are associated with medically relevant quantitative traits. The variants that we identified provide a useful starting point for studies aimed at uncovering biological mechanisms and fostering clinical translation. The power of this study to discover rare-variant associations derives from the numerous deleterious variants that are enriched in or unique to Finland. Prioritizing the sequencing of multiple population isolates that have expanded from recent bottlenecks is a strategy for increasing the scale of the discovery of rare-variant associations[7,39–41]. Because genetic drift results in a different set of alleles to pass through population-specific bottlenecks, thus enriching some variants and depleting others, the numerous rare-variant associations that could be identified by sequencing of well-phenotyped samples across multiple isolates could rapidly increase our understanding of the genetic architecture of complex traits.

Our results support recent suggestions of continuity between the genetic architectures of complex traits and disorders that are classically considered monogenic[42,43], by identifying numerous deleterious variants with large effects on quantitative traits that demonstrate geographical clustering comparable to the clustering of the mutations responsible for the Finnish Disease Heritage.

Using a Finland-specific reference panel[44] to impute FinMetSeq variants into array-genotyped samples from three other Finnish cohorts enabled us to identify additional novel associations. However, the clustering in FinMetSeq of deleterious trait-associated variants within limited geographical regions and our inability to follow up on more than 700 sub-threshold associations from FinMetSeq for which the associated variants were absent in the Finnish imputation reference panel, emphasize the importance of representing regional subpopulations in such reference panels, to account for fine-scale population structures.

The value of rare-variant studies in population isolates will depend on the richness of phenotypes in sequenced cohorts from these populations. For example, we associated fewer than 100 of the more than 24,000 deleterious, highly enriched variants identified in FinMetSeq with any of the 64 quantitative traits studied here. The associations that we identified to disease end points in FinnGen hint at the discoveries that will be possible when that database reaches its full size of 500,000 participants. The insights gained from such efforts will accelerate the implementation of precision health, informing projects in more heterogeneous populations that are still at an early stage[45].

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1457-z.

Received: 5 November 2018; Accepted: 2 July 2019;
Published online 31 July 2019.

1. Samocha, K. E. et al. Regional missense constraint improves variant deleteriousness prediction. Preprint at https://www.bioRxiv.org/content/10.1101/148353v1 (2017).
2. Marouli, E. et al. Rare and low-frequency coding variants alter human adult height. *Nature* **542**, 186–190 (2017).
3. Flannick, J. et al. Exome sequencing of 20,791 cases of type 2 diabetes and 24,440 controls. *Nature* **570**, 71–76 (2019).
4. Timpson, N. J., Greenwood, C. M. T., Soranzo, N., Lawson, D. J. & Richards, J. B. Genetic architecture: the shape of the genetic contribution to human traits and disease. *Nat. Rev. Genet.* **19**, 110–124 (2018).
5. Zuk, O. et al. Searching for missing heritability: designing rare variant association studies. *Proc. Natl Acad. Sci. USA* **111**, E455–E464 (2014).
6. Xue, Y. et al. Enrichment of low-frequency functional variants revealed by whole-genome sequencing of multiple isolated European populations. *Nat. Commun.* **8**, 15927 (2017).
7. Southam, L. et al. Whole genome sequencing and imputation in isolated populations identify genetic associations with medically-relevant complex traits. *Nat. Commun.* **8**, 15606 (2017).
8. Manolio, T. A. et al. Finding the missing heritability of complex diseases. *Nature* **461**, 747–753 (2009).
9. Jakkula, E. et al. The genome-wide patterns of variation expose significant substructure in a founder population. *Am. J. Hum. Genet.* **83**, 787–794 (2008).
10. Polvi, A. et al. The Finnish disease heritage database (FinDis) update—a database for the genes mutated in the Finnish disease heritage brought to the next-generation sequencing era. *Hum. Mutat.* **34**, 1458–1466 (2013).
11. Manning, A. et al. A low-frequency inactivating *AKT2* variant enriched in the Finnish population is associated with fasting insulin levels and type 2 diabetes risk. *Diabetes* **66**, 2019–2032 (2017).
12. Lim, E. T. et al. Distribution and medical impact of loss-of-function variants in the Finnish founder population. *PLoS Genet.* **10**, e1004494 (2014).
13. Service, S. K. et al. Re-sequencing expands our understanding of the phenotypic impact of variants at GWAS loci. *PLoS Genet.* **10**, e1004147 (2014).
14. Würtz, P. et al. Quantitative serum nuclear magnetic resonance metabolomics in large-scale epidemiology: a primer on -omic technologies. *Am. J. Epidemiol.* **186**, 1084–1096 (2017).
15. Laakso, M. et al. The Metabolic Syndrome in Men study: a resource for studies of metabolic and cardiovascular diseases. *J. Lipid Res.* **58**, 481–493 (2017).
16. Borodulin, K. et al. Forty-year trends in cardiovascular risk factors in Finland. *Eur. J. Public Health* **25**, 539–546 (2015).
17. Abraham, G. et al. Genomic prediction of coronary heart disease. *Eur. Heart J.* **37**, 3267–3278 (2016).
18. Sabatti, C. et al. Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. *Nat. Genet.* **41**, 35–46 (2009).
19. Pulizzi, N. et al. Interaction between prenatal growth and high-risk genotypes in the development of type 2 diabetes. *Diabetologia* **52**, 825–829 (2009).
20. Fagerberg, L. et al. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol. Cell. Proteomics* **13**, 397–406 (2014).
21. Corsetti, J. P. et al. Thrombospondin-4 polymorphism (A387P) predicts cardiovascular risk in postinfarction patients with high HDL cholesterol and C-reactive protein levels. *Thromb. Haemost.* **106**, 1170–1178 (2011).
22. Zhang, X. J. et al. Association between single nucleotide polymorphisms in thrombospondins genes and coronary artery disease: a meta-analysis. *Thromb. Res.* **136**, 45–51 (2015).
23. Beygo, J. et al. New insights into the imprinted MEG8-DMR in 14q32 and clinical and molecular description of novel patients with Temple syndrome. *Eur. J. Hum. Genet.* **25**, 935–945 (2017).
24. Wallace, C. et al. The imprinted DLK1-MEG3 gene region on chromosome 14q32.2 alters susceptibility to type 1 diabetes. *Nat. Genet.* **42**, 68–71 (2010).
25. Day, F. R. et al. Genomic analyses identify hundreds of variants associated with age at menarche and support a role for puberty timing in cancer risk. *Nat. Genet.* **49**, 834–841 (2017).
26. Perry, J. R. et al. Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. *Nature* **514**, 92–97 (2014).
27. Cleaton, M. A. et al. Fetus-derived DLK1 is required for maternal metabolic adaptations to pregnancy and is associated with fetal growth restriction. *Nat. Genet.* **48**, 1473–1480 (2016).
28. Chaves, J. A. et al. Genomic variation at the tips of the adaptive radiation of Darwin's finches. *Mol. Ecol.* **25**, 5282–5295 (2016).
29. Surakka, I. et al. The impact of low-frequency and rare variants on lipid levels. *Nat. Genet.* **47**, 589–597 (2015).
30. Ding, Y. et al. Plasma glycine and risk of acute myocardial infarction in patients with suspected stable angina pectoris. *J. Am. Heart Assoc.* **5**, e002621 (2015).
31. Wittemans, L. B. L. et al. Assessing the causal association of glycine with risk of cardio-metabolic diseases. *Nat. Commun.* **10**, 1060 (2019).
32. Perry, R. J. et al. Acetate mediates a microbiome–brain–β-cell axis to promote metabolic syndrome. *Nature* **534**, 213–217 (2016).
33. Tabbassum, R. et al. Genetics of human plasma lipidome: understanding lipid metabolism and its link to diseases beyond traditional lipids. Preprint at https://www.biorxiv.org/content/10.1101/457960v1 (2018).
34. Casanova, M. L. et al. Exocrine pancreatic disorders in transsgenic mice expressing human keratin 8. *J. Clin. Invest.* **103**, 1587–1595 (1999).
35. Surendran, P. et al. Trans-ancestry meta-analyses identify rare and common variants associated with blood pressure and hypertension. *Nat. Genet.* **48**, 1151–1161 (2016).
36. Liu, C. et al. Meta-analysis identifies common and rare variants influencing blood pressure and overlapping with metabolic trait loci. *Nat. Genet.* **48**, 1162–1170 (2016).
37. Palmer, C. & Pe'er, I. Statistical correction of the winner's curse explains replication variability in quantitative trait genome-wide association studies. *PLoS Genet.* **13**, e1006916 (2017).
38. Norio, R. Finnish Disease Heritage I: characteristics, causes, background. *Hum. Genet.* **112**, 441–456 (2003).
39. Service, S. et al. Magnitude and distribution of linkage disequilibrium in population isolates and implications for genome-wide association studies. *Nat. Genet.* **38**, 556–560 (2006).
40. Chiang, C. W. K. et al. Genomic history of the Sardinian population. *Nat. Genet.* **50**, 1426–1434 (2018).
41. Rivas, M. A. et al. Insights into the genetic epidemiology of Crohn's and rare diseases in the Ashkenazi Jewish population. *PLoS Genet.* **14**, e1007329 (2018).
42. Bastarache, L. et al. Phenotype risk scores identify patients with unrecognized Mendelian disease patterns. *Science* **359**, 1233–1239 (2018).
43. Niemi, M. E. K. et al. Common genetic variants contribute to risk of rare severe neurodevelopmental disorders. *Nature* **562**, 268–271 (2018).
44. Surakka, I. The rate of false polymorphisms introduced when imputing genotypes from global imputation panels. Preprint at https://www.biorxiv.org/content/10.1101/080770v1 (2016).
45. Collins, F. S. & Varmus, H. A new initiative on precision medicine. *N. Engl. J. Med.* **372**, 793–795 (2015).

METHODS

Data reporting. No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Study designs, phenotypes, and sequenced participants of the METSIM and FINRISK studies. METSIM is a single-site study investigating cardiometabolic disorders and related traits in 10,197 men randomly selected from the population register of Kuopio, Eastern Finland, aged 45 to 73 years at initial examination from 2005 to 2010. We attempted exome sequencing of all METSIM study participants[15,46].

FINRISK is a series of health examination surveys[47] based on random population samples from five (six in 2002) geographical regions of Finland, carried out every five years beginning in 1972. For exome sequencing, we chose 10,192 participants in the 1992–2007 FINRISK surveys from northeastern Finland (former provinces of North Karelia, Oulu and Lapland).

All participants in both studies provided informed consent, and study protocols were approved by the Ethics Committees at participating institutions (National Public Health Institute of Finland; Hospital District of Helsinki and Uusimaa; Hospital District of Northern Savo). All relevant ethics committees approved this study.

Selection of traits, harmonization, exclusions, covariate adjustment and transformation. Of the 257 quantitative traits measured in both METSIM and FINRISK, we selected 64 for association analysis in FinMetSeq based on clinical relevance for cardiovascular and metabolic health (Supplementary Tables 4, 5). We excluded individuals with type 1 diabetes and women who were pregnant at the time of phenotyping from all analyses; individuals with type 2 diabetes from analyses of glycaemic traits; and individuals who had not fasted for at least 8 h after their last meal for traits influenced by food consumption. A complete list of exclusions can be found in Supplementary Table 5. We adjusted measured values of systolic and diastolic blood pressures for individuals on antihypertensive medication at the time of testing[48,49], and serum lipid measures for individuals on lipid-regulating medications[50,51]. Trait adjustments are listed in Supplementary Table 5.

We prepared quantitative traits for association analysis separately for METSIM and FINRISK by linear regression on trait-specific covariates after log-transforming skewed variables. Covariates for regression analyses included: age and age^2 (METSIM); sex, age, age^2 and cohort year (FINRISK). Trait transformations and trait-specific covariates are listed in Supplementary Table 5. Several traits were adjusted for sex hormone treatment, which included women on contraceptives or hormone-replacement therapy. We transformed residuals from these initial regression analyses to normality using inverse normal scores.

Exome sequencing. We carried out exome sequencing in two phases.

Phase 1. We quantified 10,379 DNA samples with PicoGreen (ThermoFisher Scientific) and randomly parsed samples with adequate DNA (>250 ng) into cohort-specific files. We then re-arrayed samples to ensure equal numbers of METSIM and FINRISK samples on each 96-well plate, alternating samples between studies in consecutive positions within and across plates, to minimize between-study batch effects.

Using 100–250 ng input DNA, we constructed dual-indexed libraries using the HTP Library Kit (KAPA Biosystems, target insert size of 250 bp), pooling 12 libraries before hybridization to the SeqCap EZ HGSC VCRome (Roche) exome reagent. After estimating the concentration of each captured library pool by qPCR (Kapa Biosystems) to produce appropriate cluster counts for the HiSeq2000 platform (Illumina), we generated 2× 100-bp paired-end sequencing data, yielding approximately 6 Gb per sample to achieve a coverage depth of ≥20× for ≥70% of targeted bases for every sample.

Phase 2. We quantified, prepared, pooled and captured 9,937 samples as described for phase 1. We generated 2× 125-bp paired-end sequencing reads on the HiSeq2500 1T to achieve the same coverage as described for phase 1.

Contamination detection, sequence alignment, sample quality control and variant calling. We aligned sequence reads to the human genome reference build 37 (bwa-mem, v.0.7.7), realigned insertions or deletions (indels) (GATK[52] IndelRealigner v.2.4) and marked duplicates (Picard MarkDuplicates, v.1.113; http://broadinstitute.github.io/picard) and overlapping bases (BamUtil clipOverlap v.1.0.11; http://genome.sph.umich.edu/wiki/BamUtil:_clipOverlap).

For each sample, we required single-nucleotide variant (SNV) genotype array concordance >90% if SNV array data were available, excluding samples with estimated contamination >3% or sample swaps compared to existing genotype data (verifyBamID[53] v.1.1.1; Supplementary Table 1).

We called SNVs and short indels with GATK[52] (v.3.3, using recommended best practices) for all targeted exome bases and 500 bp of sequence up and downstream of each target region using HaplotypeCaller. We merged calls in batches of 200 individuals using CombineGVCFs and recalled genotypes for all individuals at all variable sites with GenotypeGVCFs.

After merging genotypes for the 19,378 samples that passed preliminary quality-control checks, we filtered SNVs and indels separately using the recommended best practices for variant quality score recalibration (VQSR). We used the true-positive variants in the GATK resource bundle (v.2.5; build37) to train the VQSR model after restricting to sites in targeted exome regions. After assessment with VQSR, we retained variants for which we identified ≥99% of true-positive sites used in the training model for both SNVs and indels.

Following initial variant filtering, we decomposed multi-allelic variants into bi-allelic variants, left-aligned indels and dropped redundant variants using vt[54] (v.0.5). We filtered variants with >2% missing calls and/or Hardy–Weinberg $P< 10^{-6}$. We additionally removed variants with an overall allele balance (alternate allele count/sum of total allele count) < 30% in genotyped samples. We excluded 86 individuals with >2% missing variant calls yielding a final analysis set of 19,292 individuals.

Array genotypes, genotype imputation and integrated exome + imputation panel. For all except 1,488 participants (57 METSIM, 1,431 FINRISK), previously generated array genotypes were available[17,55], with which we generated three datasets: (1) a merged array-based call set of all variants present in ≥90% of array-genotyped individuals across both cohorts; (2) a merged array-based Haplotype Reference Consortium (HRC) v.1.1 imputed dataset using the Michigan Imputation Server[56,57]; (3) an integrated dataset containing HRC imputed genotypes and exome-sequence variants (excluding all individuals without array data, and using the sequence-based genotypes in cases in which there was overlap between sequenced and imputed genotypes).

Annotation. We annotated the final set of sequence variants that passed quality control using variant effect predictor (VEP v.76)[58] of Ensembl using five in silico algorithms to predict the functional impact of missense variants: PolyPhen2 HumDiv and HumVar[59], LRT[60], MutationTaster[61] and SIFT[62].

Association testing. *Single variants.* We carried out single-variant association tests for transformed trait residuals with genotype dosages for variants with MAC ≥ 3 assuming an additive genetic model, using the EMMAX[63] linear mixed model approach, as implemented in EPACTS (v.3.3.0; http://genome.sph.umich.edu/wiki/EPACTS), to account for relatedness between individuals. We used genotypes for sequenced variants with MAF ≥ 1% to construct the genetic relationship matrix.

Conditioning on associated variants from previous GWAS. To differentiate association signals identified here from known associations, we performed exome-wide association analysis for each trait conditioning on variants previously associated ($P < 10^{-7}$) with that trait in the EBI GWAS catalogue (https://www.ebi.ac.uk/gwas/downloads; 4 December 2016 version)[64], publications[55,65–67] or manuscripts in preparation. The keywords from the GWAS catalogue that we used to assign known variants to each trait can be found in Supplementary Table 21. We also manually curated published associations for specific metabolites[65,68].

Using the combined HRC and exome panel, we pruned each trait-specific list of associated variants (GWAS variants) based on linkage disequilibrium ($r^2 > 0.95$). Of the 23 GWAS variants that were absent from the HRC and exome panel, we identified a proxy ($r^2 > 0.80$) variant for 17; we excluded the remaining 6 variants from the conditional analysis. The variants included in the conditional analysis are listed in Supplementary Table 22. We extracted genotypes for variants used in conditional analysis from the HRC and exome panel and converted dosages to alternate allele counts by rounding to the nearest integer (0, 1 or 2). For conditional analyses, we imputed missing genotypes for the individuals without array data using the mean genotype. We then ran association analysis using the same linear mixed model approach as in unconditional analysis but including the complete set of pruned GWAS variants as covariates in the association test. We then evaluated the novelty of conditional associations by searching OMIM, ClinVar, and the literature.

Defining loci. To identify the number of distinct associations for each trait, we performed linkage disequilibrium clumping using Swiss (https://github.com/welchr/swiss) of variants with unconditional $P < 5 \times 10^{-7}$ or both unconditional and conditional $P < 5 \times 10^{-5}$ for at least one trait. For each variant in this subset, we provided Swiss with the minimum unconditional P value across all traits. The clumping procedure starts with the variant with the smallest P value, merges into one locus all variants within ±1Mb that have $r^2 > 0.5$ with the index variant and iterates this process until no variants remain.

Calculating effects and variance explained of individual variants. For novel variants highlighted in Table 1, we evaluated the effect of each variant on the trait values by calculating the mean trait value in carriers and non-carriers. As the effect estimates from our association tests are standardized, we calculated variance explained for a given variant with the equation var. exp. $= 2f(1-f)\hat{\beta}^2$, where f is the MAF and $\hat{\beta}$ is the estimated effect size. The variance explained is included in Supplementary Table 10.

Gene-based testing. We carried out gene-based association tests using the mixed model implementation of SKAT-O[69], considering three different, but nested, sets of variants (variant 'masks'): (1) PTVs at any allele frequency with VEP annotations: frameshift_variant, initiator_codon_variant, splice_acceptor_variant, splice_donor_variant, stop_lost, stop_gained; (2) PTVs included in (1) plus

missense variants with MAF < 0.1% scored as damaging or deleterious by all five functional prediction algorithms; (3) PTVs included in (1) plus missense variants with MAF < 0.5% scored as damaging or deleterious by all five algorithms.

For each trait and mask, we only tested genes with at least two qualifying variants. Each mask contained a different number of genes with at least two qualifying variants: up to 7,996, 12,795 and 12,890 for the three masks, respectively. The exact number of genes tested varied by trait owing to sample size. We first used a Bonferroni-corrected exome-wide threshold for 12,890 genes, which corresponds to a threshold of $P < 3.88 \times 10^{-6}$. Analogous to single-variant association, we passed genes that met this association threshold for additional consideration with hierarchical false-discovery rate (FDR) correction, as described below.

Hierarchical FDR correction for testing multiple traits and variants. To control for multiple testing across 64 traits, we adopted an FDR controlling procedure[70], using a two-stage hierarchical strategy (described in the Supplementary Information). Stage 1 identifies the set of R variants (or genes) associated with at least one trait ($P < 5 \times 10^{-7}$ for single-variant unconditional results and $P < 3.88 \times 10^{-6}$ for gene-based results), controlling genome-wide FDR across all variants at $P = 0.05$. Stage 2 identifies all traits associated with the discovered variants in a manner that guarantees an average FDR $P < 0.05$.

Genotype validation. We validated exome-sequencing-based genotype calls using Sanger sequencing for METSIM carriers of 13 trait-associated very rare variants with MAF < 0.1% in seven genes, finding concordance for 107 out of 108 (99.1%) non-reference genotypes evaluated.

Replication in additional Finnish cohorts. We attempted to replicate significant single-variant associations ($P < 5 \times 10^{-7}$) and follow up suggestive single-variant associations ($P < 5 \times 10^{-5}$) using imputed array data from up to 24,776 individuals from three cohort studies: Northern Finland Birth Cohort 1966[18], the Helsinki Birth Cohort Study[19] and FINRISK study participants not included in FinMetSeq[16,17].

For each cohort, before phasing we performed genotype quality control batch-wise using standard quality thresholds. We pre-phased array genotypes with Eagle[71] (v.2.3) and imputed genotypes genome-wide with IMPUTE[72] (v.2.3.1) using 2,690 sequenced Finnish genomes and 5,092 sequenced Finnish exomes. We assessed imputation quality by confirming sex, comparing sample allele frequencies with reference population estimates and examining imputation quality (INFO score) distributions. We excluded any variant with INFO < 0.7 within a given batch from all replication/follow-up analyses.

For each cohort, we matched, harmonized, covariate adjusted and transformed available phenotypes as described above for FinMetSeq, and ran single-variant association using the EMMAX linear mixed model implemented in EPACTS, after generating kinship matrices from linkage disequilibrium-pruned (command: plink –indep-pairwise 50 5 0.2) directly genotyped variants with MAF > 5%.

Association to disease end points. From >1,100 disease end points available for analysis in FinnGen, we selected 22 that we considered most relevant to the traits analysed in FinMetSeq, identifying variant associations as described previously[33].

Association replication in UK Biobank. For eight FinMetSeq anthropometric and blood pressure traits available in UK Biobank (height, weight, body mass index, hip circumference, waist circumference, fat percentage, systolic blood pressure and diastolic blood pressure), we extracted, for variants reaching $P < 5 \times 10^{-7}$ in our combined analysis, trait-variant association statistics from http://www.nealelab.is/uk-biobank. Of the 8 traits, 7 had at least one associated variant and 23 of the total of 31 variants were available in UK Biobank. A comparison of association results is in Supplementary Table 15.

Population genetic analyses. *Identifying unrelated individuals.* To identify nearly independent common SNVs, we removed SNVs with MAF < 5% and pruned the remaining SNVs in windows of 50 SNVs, in steps of 5 SNVs, such that no pair of SNVs had $r^2 > 0.2$. We used KING[73] to estimate pairwise relationships among the exome-sequenced individuals, removing one individual from each pair inferred by KING to have a relationship of third degree or closer, yielding 14,874 unrelated individuals for population genetic analyses.

Enrichment of predicted-deleterious alleles in Finland. We assessed enrichment of predicted-deleterious alleles in Finland by comparing the 14,874 nearly unrelated FinMetSeq individuals to the 14,944 NFE control exomes in gnomAD (after removing NFE individuals from countries with substantial Finnish populations, Estonia and Sweden). We analysed the two most common alleles at each site with base quality score >10, mapping quality score >20, and coverage equal to or greater than that found in ≥80% of variable sites (17.73× in FinMetSeq, 32.27× in gnomAD), resulting in around 38.6 Mb for comparisons. We contrasted the proportional site frequency spectra for FinMetSeq and NFE for five functional variant categories (PTVs, missense, synonymous, untranslated regions and intronic variants) after down-sampling both datasets to 18,000 chromosomes.

We also assessed the enrichment of deleterious alleles within subpopulations of the FinMetSeq dataset. We applied Chromopainter and fineSTRUCTURE to 2,644 unrelated FinMetSeq individuals whose parents were both born in the same municipality to identify 16 subpopulation clusters[74] (Supplementary Information). Of the 16 clusters, we used as the reference population a cluster for which the highest proportion of the parents of its members were from early-settlement Finland (Northern Savonia population 3 (NSv3), Supplementary Table 17). We used the twelve clusters with >100 members in subsequent analyses (Supplementary Table 17). We then compared the ratio of the site frequency spectra to the reference for PTVs, missense and synonymous variants, down-sampling both datasets to 200 haploid chromosomes. For each comparison, we computed statistical evidence for enrichment or depletion at a given allele count bin by exact binomial test against a null of equal number of variants found in both the test and reference cluster.

Geographical clustering of predicted functionally deleterious alleles. We first generated a distance matrix tabulating the pairwise geographical distance between the birthplaces of all available parents of unrelated sequenced individuals. For each variant of interest, we computed for the minor allele carriers in FinMetSeq the mean distance among all parent pairs. We evaluated statistical significance of geographical clustering by comparing the observed mean distance to mean distances for up to 10,000,000 sets of randomly drawn non-carrier individuals matched by cohort status and number of parents with birthplace information available.

To assess whether PTVs or missense variants may be more geographically clustered than synonymous variants, we first identified a set of near-independent variants ($r^2 > 0.02$) with MAC ≥ 3 and MAF ≤ 5% among the 14,874 unrelated individuals. For each variant, we computed the mean pairwise geographical distance between the birthplaces across all pairs of the available parents of carriers of the minor allele and regressed this mean distance on variant class (PTVs, missense or synonymous) and MAC, MAC^2 and MAC^3 (Supplementary Table 16). For those variants in gnomAD, we also assessed whether variants enriched in FinMetSeq compared to NFE are more likely to be geographically clustered. As above, we computed the mean pairwise distances among parents of carriers of the minor allele and regressed mean distance on the logarithm of enrichment and MAC, MAC^2 and MAC^3 (Supplementary Table 19). In both analyses, we assessed a model with the interaction terms but report only the model without interactions if the interactions were not significant.

Heritability estimates and genetic correlations. We used genome-wide array genotype data on the 13,326 unrelated individuals for whom both exome sequencing and array data were available to estimate heritability and genetic correlations for the 64 traits. We constructed a genetic relationship matrix with PLINK[75] (v.1.90b, https://www.cog-genomics.org/plink2) by applying additional filters for MAF > 1% and genotype missingness rate < 2% to the set of previously used genotyped SNVs, leaving 205,149 SNVs for genetic relationship matrix calculation. We used the exact mixed model approach of biMM[76] (v.1.0.0, http://www.helsinki.fi/~mjxpirin/download.html) to estimate the heritability of our 64 traits and the genetic correlation of the 2,016 trait pairs.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The sequencing data can be accessed through dbGaP (https://www.ncbi.nlm.nih.gov/gap/) using study numbers phs000756 and phs000752. Association results can be accessed at http://pheweb.sph.umich.edu/FinMetSeq/ and are searchable via the Type 2 Diabetes Knowledge Portal (http://www.type2diabetesgenetics.org/). Summary statistics are also available through the NHGRI-EBI GWAS Catalog at https://www.ebi.ac.uk/gwas/downloads/summary-statistics.

46. Stancáková, A. et al. Changes in insulin sensitivity and insulin release in relation to glycemia and glucose tolerance in 6,414 Finnish men. *Diabetes* **58**, 1212–1221 (2009).
47. Borodulin, K. et al. Cohort profile: the National FINRISK Study. *Int. J. Epidemiol.* **47**, 696–696i (2017).
48. Wu, J. et al. A summary of the effects of antihypertensive medications on measured blood pressure. *Am. J. Hypertens.* **18**, 935–942 (2005).
49. Tobin, M. D., Sheehan, N. A., Scurrah, K. J. & Burton, P. R. Adjusting for treatment effects in studies of quantitative traits: antihypertensive therapy and systolic blood pressure. *Stat. Med.* **24**, 2911–2935 (2005).
50. Liu, D. J. et al. Exome-wide association study of plasma lipids in >300,000 individuals. *Nat. Genet.* **49**, 1758–1766 (2017).
51. Friedewald, W. T., Levy, R. I. & Fredrickson, D. S. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin. Chem.* **18**, 499–502 (1972).
52. DePristo, M. A. et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011).
53. Jun, G. et al. Detecting and estimating contamination of human DNA samples in sequencing and array-based genotype data. *Am. J. Hum. Genet.* **91**, 839–848 (2012).
54. Tan, A., Abecasis, G. R. & Kang, H. M. Unified representation of genetic variants. *Bioinformatics* **31**, 2202–2204 (2015).

55. Davis, J. P. et al. Common, low-frequency, and rare genetic variants associated with lipoprotein subclasses and triglyceride measures in Finnish men from the METSIM study. *PLoS Genet.* **13**, e1007079 (2017).
56. Das, S. et al. Next-generation genotype imputation service and methods. *Nat. Genet.* **48**, 1284–1287 (2016).
57. The Haplotype Reference Consortium. A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* **48**, 1279–1283 (2016).
58. McLaren, W. et al. The Ensembl Variant Effect Predictor. *Genome Biol.* **17**, 122 (2016).
59. Adzhubei, I. A. et al. A method and server for predicting damaging missense mutations. *Nat. Methods* **7**, 248–249 (2010).
60. Chun, S. & Fay, J. C. Identification of deleterious mutations within three human genomes. *Genome Res.* **19**, 1553–1561 (2009).
61. Schwarz, J. M., Cooper, D. N., Schuelke, M. & Seelow, D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat. Methods* **11**, 361–362 (2014).
62. Kumar, P., Henikoff, S. & Ng, P. C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* **4**, 1073–1081 (2009).
63. Kang, H. M. et al. Variance component model to account for sample structure in genome-wide association studies. *Nat. Genet.* **42**, 348–354 (2010).
64. Buniello, A. et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* **47**, D1005–D1012 (2019).
65. Kettunen, J. et al. Genome-wide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA. *Nat. Commun.* **7**, 11122 (2016).
66. Kettunen, J. et al. Genome-wide association study identifies multiple loci influencing human serum metabolite levels. *Nat. Genet.* **44**, 269–276 (2012).
67. Teslovich, T. M. et al. Identification of seven novel loci associated with amino acid levels using single-variant and gene-based tests in 8545 Finnish men from the METSIM study. *Hum. Mol. Genet.* **27**, 1664–1674 (2018).
68. Inouye, M. et al. Novel loci for metabolic networks and multi-tissue expression studies reveal genes for atherosclerosis. *PLoS Genet.* **8**, e1002907 (2012).
69. Lee, S. et al. Optimal unified approach for rare-variant association testing with application to small-sample case–control whole-exome sequencing studies. *Am. J. Hum. Genet.* **91**, 224–237 (2012).
70. Peterson, C. B., Bogomolov, M., Benjamini, Y. & Sabatti, C. Many phenotypes without many false discoveries: error controlling strategies for multitrait association studies. *Genet. Epidemiol.* **40**, 45–56 (2016).
71. Loh, P. R. et al. Reference-based phasing using the Haplotype Reference Consortium panel. *Nat. Genet.* **48**, 1443–1448 (2016).
72. Howie, B. N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* **5**, e1000529 (2009).
73. Manichaikul, A. et al. Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**, 2867–2873 (2010).
74. Lawson, D. J., Hellenthal, G., Myers, S. & Falush, D. Inference of population structure using dense haplotype data. *PLoS Genet.* **8**, e1002453 (2012).
75. Chang, C. C. et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
76. Pirinen, M. et al. biMM: efficient estimation of genetic variances and covariances for cohorts with high-dimensional phenotype measurements. *Bioinformatics* **33**, 2405–2407 (2017).

Acknowledgements We thank T. Teshiba for coordinating ethical permissions and samples; S. Kerminen, D. Lawson and G. Busby for discussions and providing scripts to run fineSTRUCTURE. S.R. was supported by the Academy of Finland Center of Excellence in Complex Disease Genetics (312062), Academy of Finland (285380), the Finnish Foundation for Cardiovascular Research, the Sigrid Juselius Foundation, Biocentrum Helsinki and University of Helsinki HiLIFE Fellow grant. V.R. acknowledges support by RFBR, research project 18-04-00789 A. V.S. was supported by the Finnish Foundation for Cardiovascular Research. C.S. and L.S. received funding from HG006695, HL113315 and MH105578. M.A.-K. is supported by a Senior Research Fellowship from the National Health and Medical Research Council (NHMRC) of Australia (APP1158958) and works in a unit that is supported by the University of Bristol and UK Medical Research Council (MC_UU_12013/1). The Baker Institute is supported in part by the Victorian Government's Operational Infrastructure Support Program. A.U.J., D.R., L.J.S., H.M.S., R.W., P.Y., X.Y. and M.B. received funding from DK062370. S.K.S., C.W.K.C. and N.B.F. received funding from HL113315 and NS062691. The METSIM study was supported by grants from Academy of Finland (321428), the Sigrid Juselius Foundation, the Finnish Foundation for Cardiovascular Research, Kuopio University Hospital and the Centre of Excellence of Cardiovascular and Metabolic Diseases is supported by the Academy of Finland (M.L.). Sequencing was funded by 5U54HG003079. A.E.L., K.M.S., H.J.A., C.C.C., C.J.K., K.L.K., D.C.K., D.E.L., J.N., T.J.N., S.K.D., N.O.S., I.M.H. and R.K.W. were funded by 5U54HG003079 and 5UM1HG008853-03.

Author contributions A.E.L., L.J.S., R.K.W., A. Palotie, V.S., M.L., S.R., M.B. and N.B.F. designed the study. A.E.L., K.M.S., H.J.A., R.S.F., D.C.K., D.E.L., J.N., T.J.N. and J.V. produced and quality-controlled the sequence data. A.E.L., A.S.H., A.U.J., A. Pietilä, H.M.S., M.A.-K., V.S. and M.L. collected, quality-controlled and/or prepared the clinical data for association analysis. A.E.L., K.M.S., C.W.K.C., S.K.S., A.S.H., L.S., M.P., C.C.C., A.U.J., C.J.K., K.L.K., V.R., D.R., J.V., R.W., P.Y. and X.Y. analysed data. A.S.H., J.G.E., M.A.-K., M.-R.J. and M.M. collected, quality-controlled and analysed replication data. H.L., S.K.D., N.O.S., I.M.H., C.S., S.R., M.B. and N.B.F. supervised experiments and analyses. A.E.L., K.M.S., C.W.K.C., S.K.S., C.S., M.B. and N.B.F. wrote the paper.

Competing interests : V.S. has participated in a conference trip sponsored by Novo Nordisk and received a honorarium from the same source for participating in an advisory board meeting. He also has ongoing research collaboration with Bayer. H.L. is a member of the Nordic Expert group unconditionally supported by Gedeon Richter Nordics and has received an honorarium from Orion. All other authors have no competing interests.

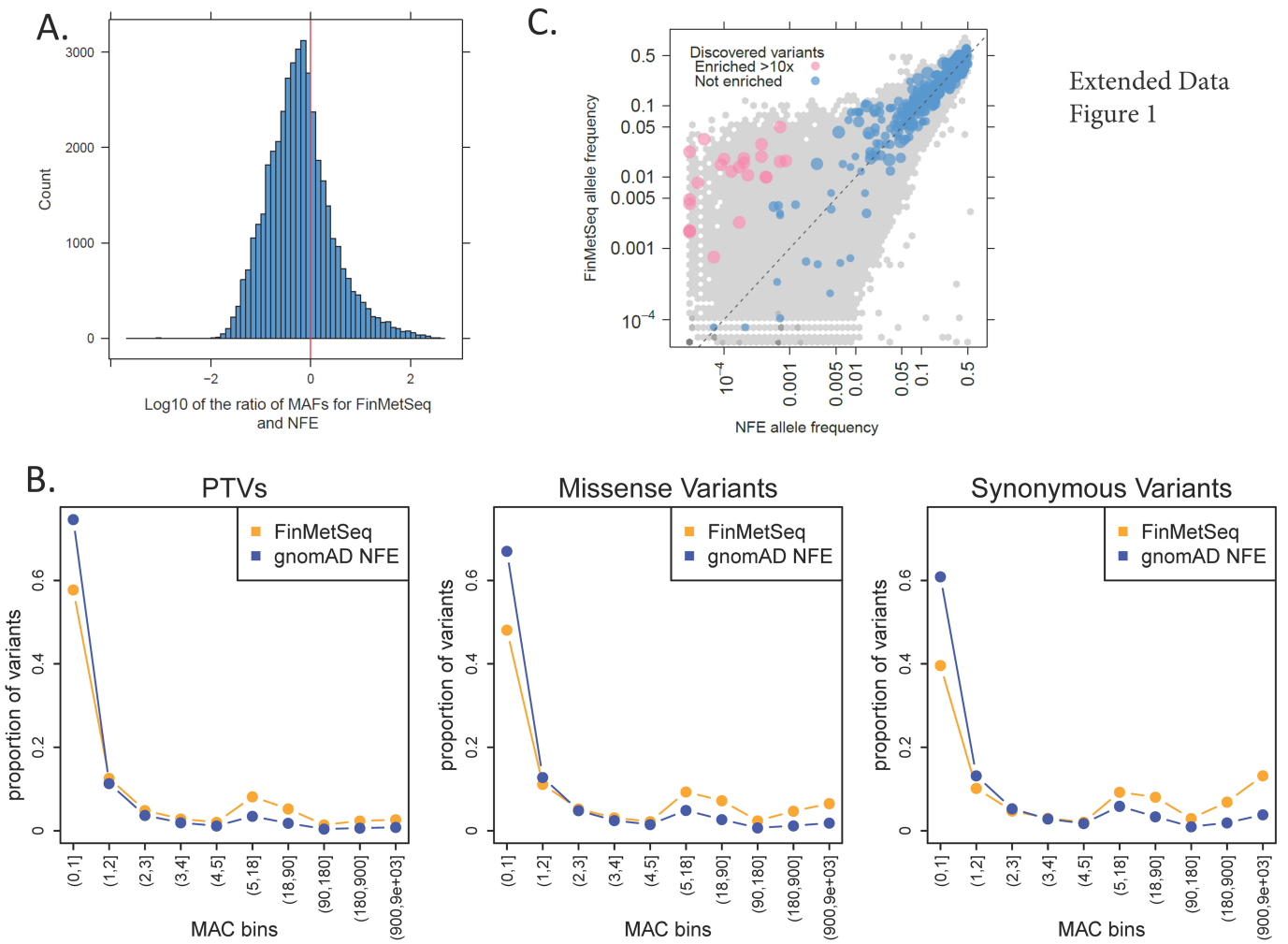
Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1457-z.

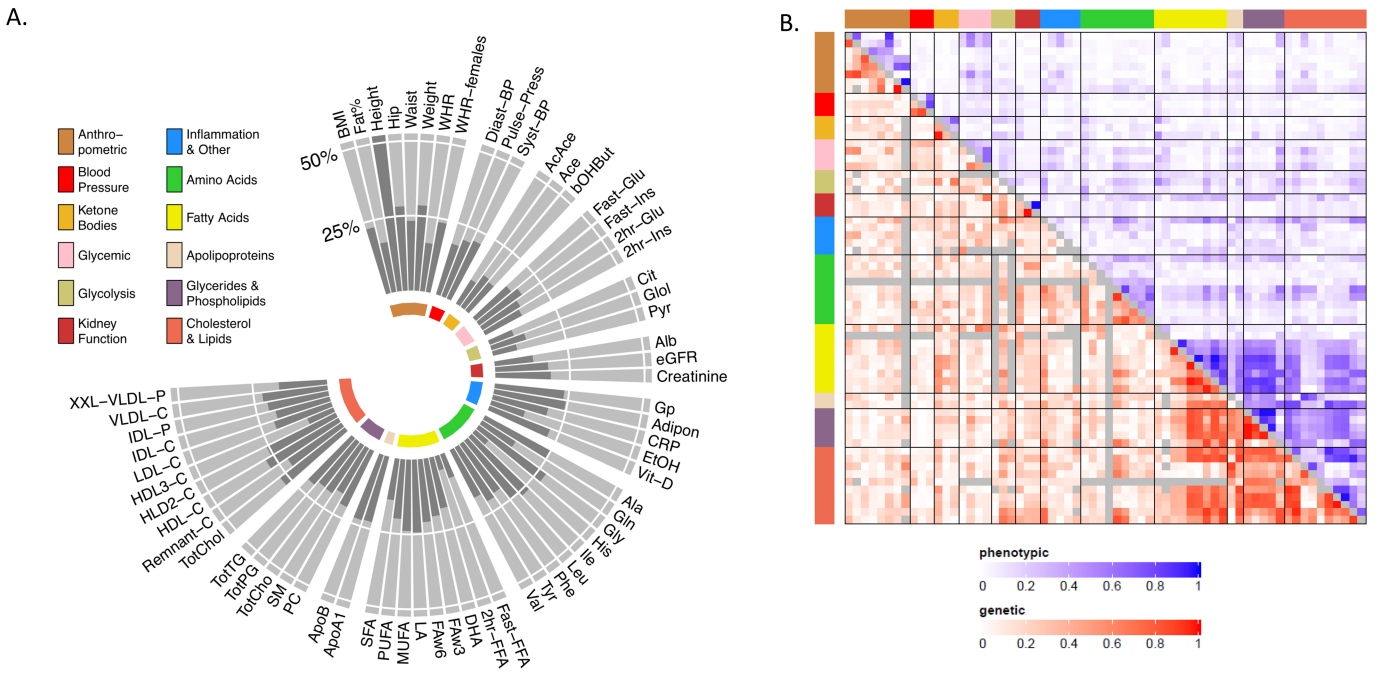
Correspondence and requests for materials should be addressed to M.B. or N.B.F.

Peer review information *Nature* thanks Timothy Frayling, Alan Shuldiner, André G. Uitterlinden, Daniel E. Weeks for their contribution to the peer review of this work.

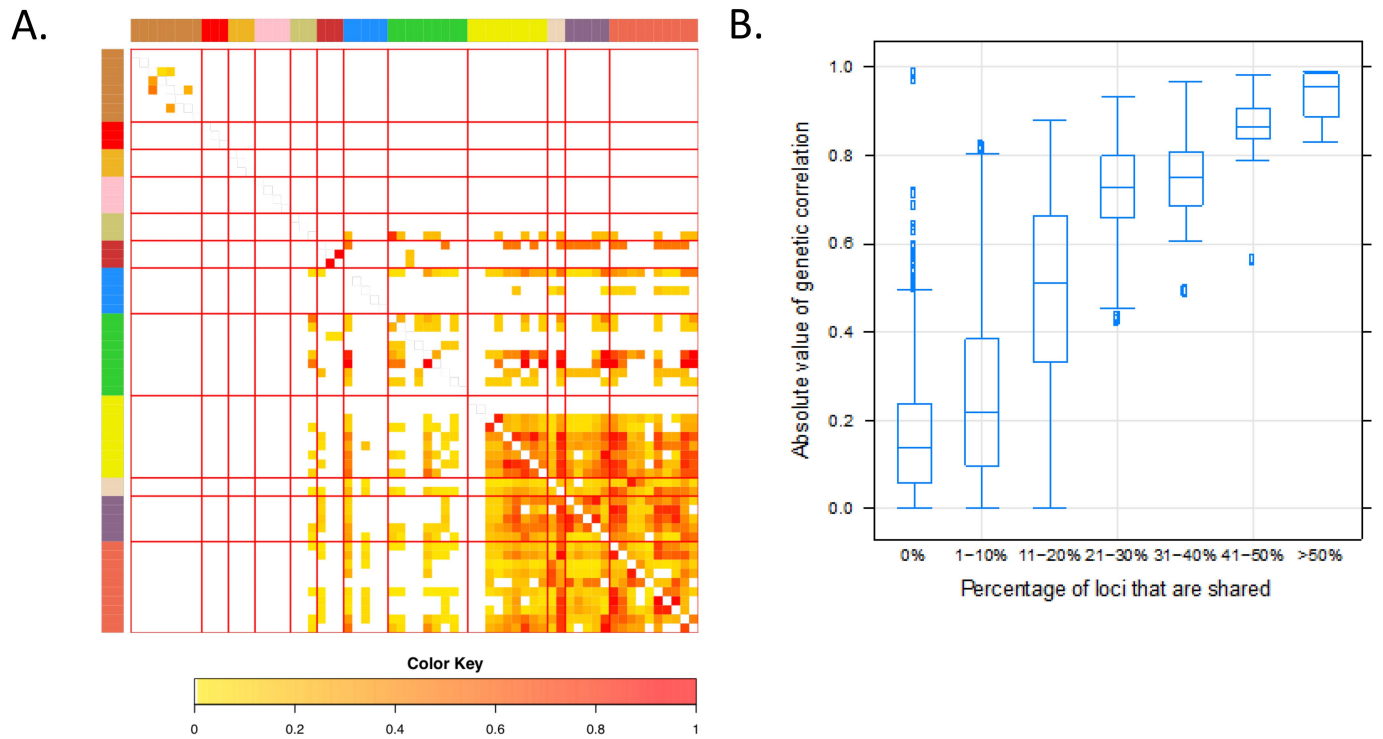
Extended Data Figure 1



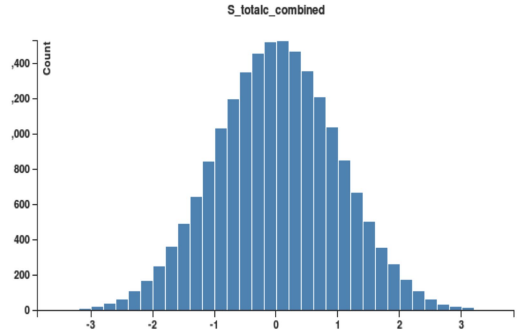
Extended Data Fig. 1 | Allele frequency comparisons between FinMetSeq and NFE from gnomAD. a, Distribution of allelic frequencies between FinMetSeq and gnomAD NFE. The comparison of allele frequencies shows the excess of variants at higher frequency in Finland as a result of the multiple bottlenecks experienced in Finnish population history. **b**, Proportional site frequency spectra between FinMetSeq and gnomAD NFE by variant annotation class. In general, we find a depletion of the variants in the rarest frequency class, as well as enrichment of variants in the intermediate to common frequency range. The site frequency spectra were down-sampled to 18,000 chromosomes for each data set. **c**, Comparison of MAFs for trait-associated variants in FinMetSeq and NFE gnomAD. Plotted in the grey background is a two-dimensional histogram of variants with non-zero allele frequencies in both gnomAD and FinMetSeq but no trait associations. Variants associated with at least one trait are coloured and scaled inversely proportional to the logarithm of the association *P* value. Variants >10× enriched in FinMetSeq compared to NFE are pink, those <10× enriched are in blue. The dashed line is the line of equal frequency. Two-sided uncorrected *P* values are from a regression of trait on the count of alternative allele at each variant. The number of independent individuals used in each point is listed in Supplementary Table 5.



Extended Data Fig. 2 | Heritability of and correlations between traits. **a**, **b**, Traits are in the same order, clockwise in **a**, and left to right and top to bottom in **b**, following the trait group colour key. **a**, Heritability estimated in 13,342 unrelated individuals (for abbreviations see Supplementary Table 4; for details see Supplementary Table 6). **b**, Heat map of the absolute Pearson correlations of standardized trait values (top right triangle) and the absolute values of estimated pairwise genetic correlations (bottom left triangle). Genetic correlations are estimated in 13,342 unrelated individuals. Values in grey below the diagonal had trait heritability less than 1.5× the s.e. of heritability.



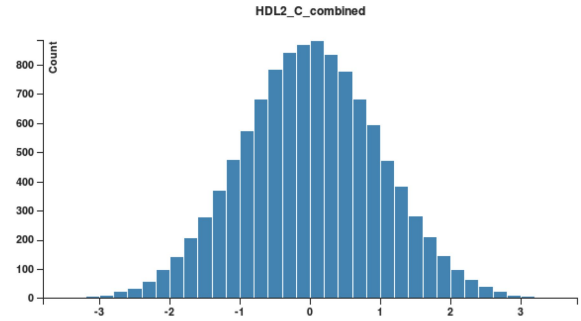
Extended Data Fig. 3 | Properties of associations shared between traits. **a**, Shared genomic associations by pairs of traits. For traits *x* and *y*, colour in row *x* and column *y* reflects the number of loci associated with both traits divided by the number of loci associated with trait *x*. Traits are presented in the same order as in Extended Data Fig. 2a, and the side and top colour bars reflect trait groups. **b**, Relationship between estimated genetic correlation and extent of sharing of genetic associations. For each trait pair, the extent of locus sharing is defined as the number of loci associated with both traits divided by the total number of loci associated with either trait. Analysis using the absolute value of the Pearson correlation of the residual series results in a very similar pattern. The number of trait pairs in each *x*-axis category is as follows: 0–1%, 819; 1–10%, 204; 11–20%, 102; 21–30%, 41; 31–40%, 29; 41–50%, 16; >50%, 13. The bar within each box is the median, the box represents the upper and lower quartiles, whiskers extend to 1.5× the interquartile range and points represent outliers.



Symbol	Meaning
▲	Rare homozygote
▼	Heterozygote
⋮	Mean trait value (all samples)
|	Mean trait value (variant carriers)

GENE	mmskat.P	VARIANT	SV.P	BETA	MAF	MAC
APOB	2.0e-13	2:21228642_G/A	0.93	0.089	0.000030	1.0
		2:21230094_AT/A	0.0000013	-2.8	0.000080	3.0
		2:21230336_AT/A	0.33	-0.96	0.000030	1.0
		2:21230583_C/A	0.00011	-2.7	0.000050	2.0
		2:21231852_GA/G	0.0065	-2.7	0.000030	1.0
		2:21233909_TGA/T	0.16	-1.4	0.000030	1.0
		2:21234140_C/CG	0.052	-1.9	0.000030	1.0
		2:21234858_C/CG	0.0098	-2.5	0.000030	1.0
		2:21251412_T/C	0.23	-1.2	0.000030	1.0

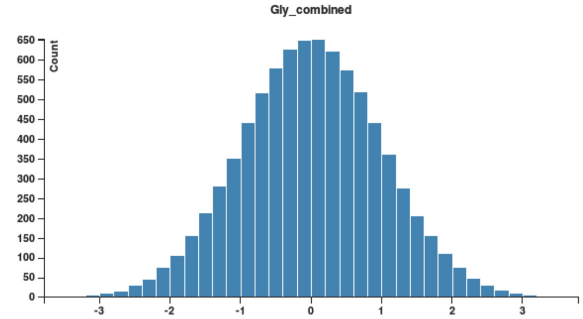
Extended Data Fig. 4 | Gene-based association of extremely rare variants in *APOB* with serum total cholesterol. Top, the distribution of the covariate-adjusted and inverse-normal transformed phenotype. Bottom, the association statistics for each variant included in the gene-based test along with the trait value for minor allele carriers of each variant (orange triangles). SV.P is the *P* value from the analysis of each variant in a single-variant analysis. The number of independent individuals in the analysis is 19,291.



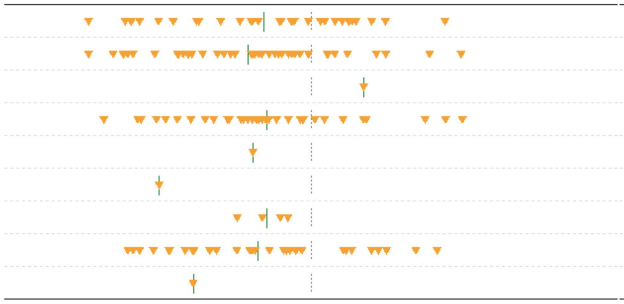
Symbol	Meaning
▲	Rare homozygote
▼	Heterozygote
⋮	Mean trait value (all samples)
|	Mean trait value (variant carriers)

GENE	mmskat.P	VARIANT	SV.P	BETA	MAF	MAC
SECTM1	5.6e-7	17:80280768_G/A	0.0088	-2.6	0.000050	1.0
		17:80280876_G/A	0.000020	-2.9	0.000090	2.0

Extended Data Fig. 5 | Gene-based association of rare variants in *SECTM1* with HDL2 cholesterol. Top, the distribution of the covariate-adjusted and inverse-normal transformed phenotype. Bottom, the association statistics for each variant included in the gene-based test, along with the trait value for minor allele carriers of each variant (orange triangles). SV.P is the *P* value from the analysis of each variant in a single-variant analysis. The number of independent individuals in the analysis is 10,984.

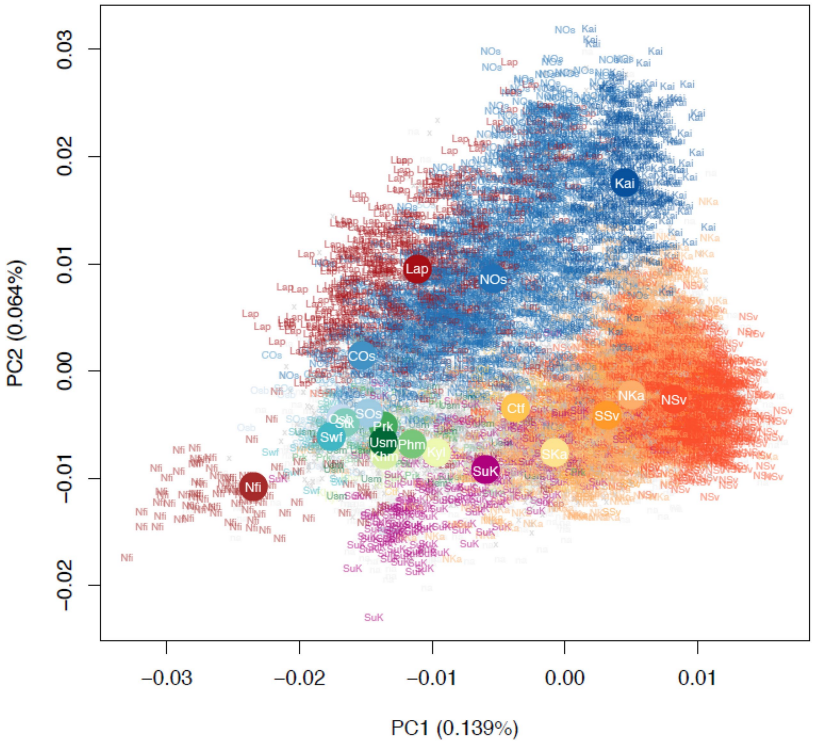


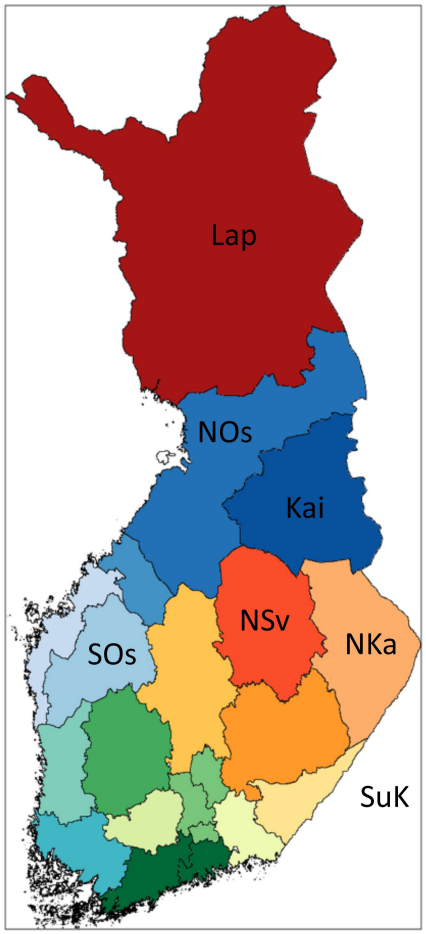
Symbol	Meaning
▲	Rare homozygote
▼	Heterozygote
⋮	Mean trait value (all samples)
|	Mean trait value (variant carriers)



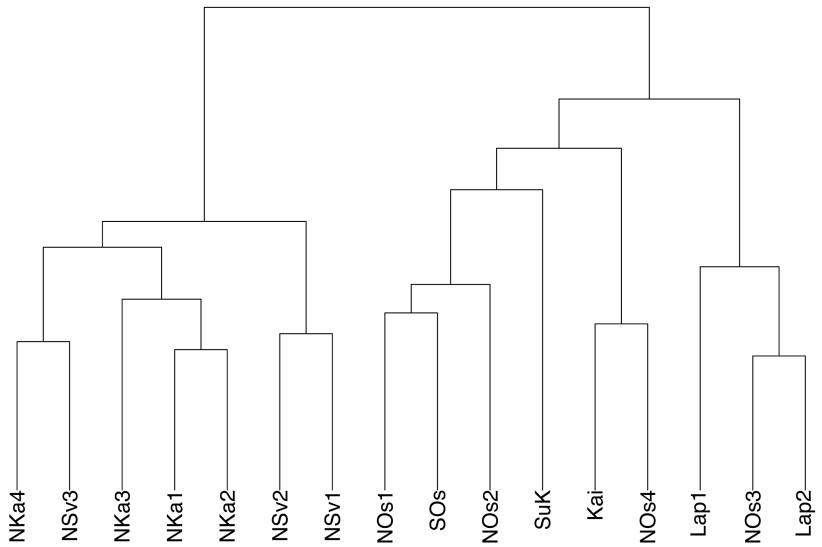
GENE	mmskat.P	VARIANT	SV.P	BETA	MAF	MAC
ALDH1L1	3.9e-21	3:125828918_G/A	0.000029	-0.73	0.0020	33
		3:125831672_G/A	1.8e-8	-0.87	0.0026	42
		3:125850306_T/C	0.53	0.62	0.000060	1.0
		3:125854377_C/T	0.000072	-0.64	0.0024	39
		3:125873443_C/T	0.44	-0.76	0.000060	1.0
		3:125877290_G/A	0.039	-2.1	0.000060	1.0
		3:125879703_G/T	0.19	-0.66	0.00024	4.0
		3:125879755_C/T	0.000050	-0.70	0.0021	34
		3:125879807_T/A	0.13	-1.5	0.000060	1.0

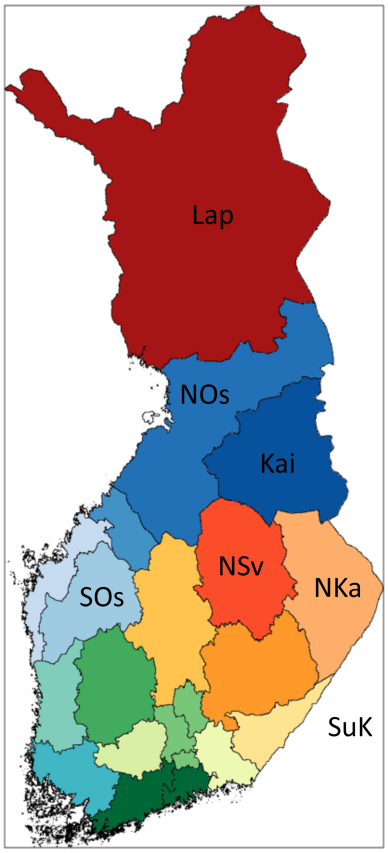
Extended Data Fig. 6 | Gene-based association of extremely rare variants in *ALDH1L1* with glycine levels. Top, the distribution of the covariate-adjusted and inverse-normal transformed phenotype. Bottom, the association statistics for each variant included in the gene-based test, along with the trait value for minor allele carriers of each variant (orange triangles). SV.P is the *P* value from the analysis of each variant in a single-variant analysis. The number of independent individuals in the analysis is 8,206.



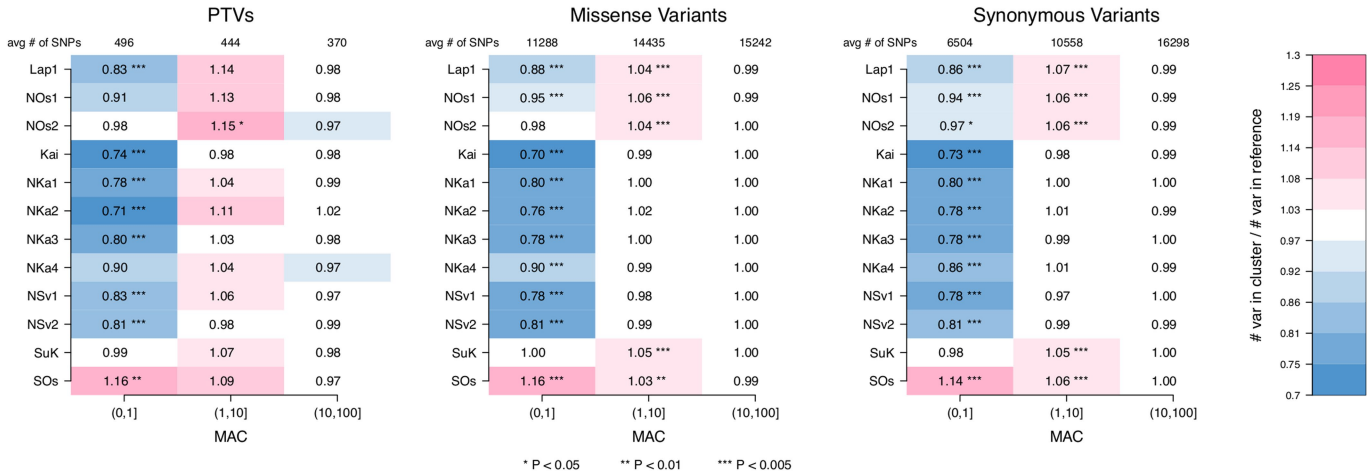


Extended Data Fig. 7 | Population structure of the FinMetSeq dataset, by region. Population structure, by region, from a principal component analysis of exome-sequencing variant data (MAF > 1%) for 14,874 unrelated individuals with known parental birthplaces. Colour indicates individuals with both parents born in the same region; grey indicates individuals with different parental birth regions or missing information for one parent. Ctf, Central Finland; COs, Central Ostrobothnia; Kai, Kainuu; Khm, Kanta-Hame; Kyl, Kymenlaakso; Lap, Lapland; Nka, Northern Karelia; NOs, Northern Ostrobothnia; NSv, Northern Savonia; Osb, Ostrobothnia; Phm, Paijat-Hame; Prk, Pirkanmaa; SKa, Southern Karelia; SOs, Southern Ostrobothnia; SSv, Southern Savonia; Stk, Satakunta; Swf, Southwest Finland; Usm, Uusimaa; X, split parental birthplaces. Large solid circles represent the centre of each region. A map of Finland with regions labelled is supplied for reference.





Extended Data Fig. 8 | Hierarchical clustering tree produced by fineSTRUCTURE. We identified 16 subpopulations within the FinMetSeq dataset by applying a haplotype-based clustering algorithm, fineSTRUCTURE, on 2,644 unrelated individuals born by 1955 whose parents were both born in the same municipality (Methods). Each subpopulation is named based on the most common parental birth location among its members. Kai, Kainuu; Lap, Lapland; NKa, North Karelia; NOs, North Ostrobothnia; NSv, North Savonia; SOs, South Ostrobothnia; SuK, Surrendered Karelia. A map of Finland with regions labelled is supplied for reference. If multiple subpopulations share the same location label, the subpopulation is further distinguished with a numeral. NSv3 is used as an internal reference for the enrichment analysis. See Supplementary Table 17 for more detailed demographic descriptions of each subpopulation.



Extended Data Fig. 9 | Regional variation in allele frequencies by functional annotation. Enrichment of variants by allelic class in regional subpopulations of late-settlement Finland (defined in Supplementary Table 17). Each bin represents the ratio of variants in the subpopulation compared to the reference subpopulation (NSv3), after down-sampling the frequency spectra of all populations to 200 chromosomes. Pink cells represent enrichment (ratio >1), blue cells represent depletion (ratio <1). Sample sizes and confidence intervals for each enrichment ratio and the associated *P* values are presented in Supplementary Table 18. The results are consistent with multiple bottlenecks in late-settlement Finland, particularly for populations in Lapland and Northern Ostrobothnia. $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.005$.

nature research

Corresponding author(s): Freimer, NB

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	bwa-mem v0.7.7 picard v1.113 GATK - IndelRealigner v2.4 BamUtil - clipOverlap v1.0.11 verifyBamID v1.1.1 GATK v3.3 VQSR vt v0.5 VEP v76 CADD v1.2 (PolyPhen2 (v2.2.2), LRT (11/09 release), MutationTaster (2013 release), SIFT (09/11 release)) as in dbNSFP v2.4
Data analysis	EMMAX EPACTS v3.3.0 PLINK v1.9 R 3.4.0 Swiss v1.0.0 SKAT-O

Eagle v2.3
IMPUTE2 v2.3.1
KING v2.0
SHAPEIT v2, r837
ChromoPainter v2
fineStructure v2.0.8
biMM v1.0.0

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The sequence data can be accessed through dbGaP using the following study numbers: FINRISK: phs000756, METSIM: phs000752. Association results can be accessed at http://pheweb.sph.umich.edu/FinMetSeq/. NOTE: METSIM phs000752 is the correct accession number, however dbGaP has not yet released the data. We are working to resolve this

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	All available subjects in two extensive population cohorts of Finnish subjects
Data exclusions	We excluded 126 individuals, 92 with type 1 diabetes and 34 women who were pregnant at the time of phenotyping, from all analyses. Pregnancy is known to dramatically alter metabolic profiles and type 1 diabetics also represent an altered profile compared to the general population, and thus both might obscure variant-trait relationships present in the rest of the population. Both represent a very small fraction of the overall sample. Though these samples were sequenced, they were excluded prior to any gene/trait association testing. We also excluded 3,088 individuals with T2D from analyses of glycemic traits. For traits influenced by food consumption (amino acids, fatty acids, LDL cholesterol, total triglycerides, and glycemic traits), we excluded individuals not fasting for at least 8 hours after their last meal. A complete list of exclusions can be found in Supplementary Table 4. All exclusion criteria were determined before any analyses were conducted.
Replication	We performed replication analysis of significant single-variant associations (P<5×10-7) and follow-up analysis of suggestive single-variant associations (P<5×10-5) in up to 24,776 individuals from three GWAS cohort studies: Northern Finland Birth Cohort 1966 (NFBC1966), the Helsinki Birth Cohort Study (HBCS), and FINRISK study participants not included in the exome sequencing portion of FinMetSeq. We also did look ups of our discoveries in UK Bio Bank (for some of the same quantitative traits) and FinnGen (a Finnish Biobank, for disease endpoints).
Randomization	no experimental treatments in our study
Blinding	no experimental treatments in our study

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	METSIM is a single-site study comprised of 10,197 men randomly selected from the population register of Kuopio, Eastern Finland, aged 45 to 73 years at initial examination from 2005 to 2010. FINRISK is a series of health examination surveys carried out by the National Institute for Health and Welfare (formerly National Public Health Institute) of Finland every five years beginning in 1972. The surveys are based on random population samples from five (six in 2002) geographical regions of Finland. Participants were selected by 10-year age group, sex, and study area. Survey sample sizes have varied from 7,000 to 13,000 individuals and participation rates from 60% to 90%. The age-range was 25 to 64 years until 1992 and 25 to 74 years since 1997.
Recruitment	FINRISK - Multi-site national health examination of adults executed every 5 years since 1972 representing a geographically diverse cross-section of the country. No major exclusions. METSIM - Single site population cohort representing older (>= 45 at recruitment) adult males in the city of Kuopio in eastern Finland. Though a population cohort, recruited only older men due to their increased risk for cardiovascular and metabolic disease.

ARTICLE

https://doi.org/10.1038/s41586-019-1451-5

Human placenta has no microbiome but can contain potential pathogens

Marcus C. de Goffau[1,2,8], Susanne Lager[3,4,5,8], Ulla Sovio[3,4], Francesca Gaccioli[3,4], Emma Cook[3], Sharon J. Peacock[1,6,7], Julian Parkhill[1,2]*, D. Stephen Charnock-Jones[3,4,9] & Gordon C. S. Smith[3,4,9]*

We sought to determine whether pre-eclampsia, spontaneous preterm birth or the delivery of infants who are small for gestational age were associated with the presence of bacterial DNA in the human placenta. Here we show that there was no evidence for the presence of bacteria in the large majority of placental samples, from both complicated and uncomplicated pregnancies. Almost all signals were related either to the acquisition of bacteria during labour and delivery, or to contamination of laboratory reagents with bacterial DNA. The exception was *Streptococcus agalactiae* (group B Streptococcus), for which non-contaminant signals were detected in approximately 5% of samples collected before the onset of labour. We conclude that bacterial infection of the placenta is not a common cause of adverse pregnancy outcome and that the human placenta does not have a microbiome, but it does represent a potential site of perinatal acquisition of *S. agalactiae*, a major cause of neonatal sepsis.

Placental dysfunction is associated with common adverse pregnancy outcomes that determine a substantial proportion of the global burden of disease[1]. However, the cause of placental dysfunction in most cases is unknown. Several studies have used sequencing-based methods for bacterial detection (metagenomics and 16S rRNA gene amplicon sequencing), and have concluded that the placenta is physiologically colonized by a diverse population of bacteria (the 'placental microbiome') and that the nature of this colonization may differ between healthy and complicated pregnancies[2–4]. This contrasts with the view in the pre-sequencing era that the placenta was normally sterile[5]. However, several studies that applied sequencing-based methods informed by the potential for false-positive results due to contamination[6–8] have failed to detect a placental microbiome[9–12]. The aim of the present study was to determine whether pre-eclampsia, delivery of a small for gestational age (SGA) infant and spontaneous preterm birth (PTB) were associated with the presence or a pattern of bacterial DNA in the placenta and to determine whether there was evidence to support the existence of a placental microbiome. We used samples from a large, prospective cohort study of nulliparous pregnant women[13], and applied an experimental approach informed by the potential for false-positive results[14].

Experimental approach

We studied two cohorts of patients (Extended Data Fig. 1 and Supplementary Tables 1, 2). In cohort 1, babies were all delivered by pre-labour Caesarean section, and the cohort included 20 patients with pre-eclampsia, 20 SGA infants, and 40 matched controls. The placental biopsies were spiked with approximately 1,100 colony-forming units (CFUs) of *Salmonella bongori* (positive control) and samples were analysed using both deep metagenomic sequencing of total DNA (424 million reads on average per sample) and 16S rRNA gene amplicon sequencing. Cohort 2 included 100 patients with pre-eclampsia, 100 SGA infants, 198 matched controls (two controls were used twice) and 100 preterm births. All of these samples were analysed twice using 16S rRNA gene amplicon sequencing from DNA extracted by two different kits.

Cohort 1: metagenomics and 16S rRNA

The positive control (*S. bongori*, average 180 reads per sample, Extended Data Fig. 2a) was detected in all samples. Several other bacterial signals were also observed. Principal component analysis (PCA) (Fig. 1a) demonstrated that almost all of the variation in the metagenomics data (98%) was represented by principal components 1 (80%) and 2 (18%). This variation was driven by batch effects and not by case–control status (Fig. 1b). Any variation that is associated with processing batches, and not the sampling framework, must be due to contamination. A heat map (Fig. 1c) showed that eight out of the ten runs had a pronounced *Escherichia coli* signal (more than 20,000 reads in 64 samples, and 50–150 reads in 16 samples), a large collection of additional bacterial signals, and high levels of PhiX174 reads (group 1; Fig. 1c). Additional analyses mapping all *E. coli* reads from all samples together against the closest reference genome (WG5) showed that all *E. coli* reads belonged to the same strain (Extended Data Fig. 3) and are, therefore, due to contamination. All samples belonging to runs 4 and 5 (Fig. 1b) also had strong *Bradyrhizobium* and *Rhodopseudomonas palustris* signals (group 2 in PCA analysis). Runs 8 and 9 (group 3) lacked these strong signals. Two samples had strong human betaherpesvirus 6B (HHV-6B) signals (more than 10,000 read pairs; Fig. 1a–c), which reflected inheritance of the chromosomally integrated virus, affecting 0.5–1% of individuals in western populations[15].

We analysed the concordance between metagenomics and 16S rRNA gene amplicon sequencing in 79 samples from cohort 1 (Table 1, one 16S primer pair failed). The only signal consistently detected using both methods was *S. bongori*. An average of approximately 33,000 *S. bongori* reads (54% of total reads) were found by 16S rRNA amplicon sequencing (Extended Data Fig. 2b). *S. bongori* was not detected in the 16S negative controls (DNA extraction blanks; Table 1). The level of agreement between metagenomics and 16S rRNA for the other

[1]Wellcome Sanger Institute, Cambridge, UK. [2]Department of Veterinary Medicine, University of Cambridge, Cambridge, UK. [3]Department of Obstetrics and Gynaecology, University of Cambridge, National Institute for Health Research Biomedical Research Centre, Cambridge, UK. [4]Centre for Trophoblast Research (CTR), Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, UK. [5]Department of Women's and Children's Health, Uppsala University, Uppsala, Sweden. [6]Department of Medicine, University of Cambridge, Cambridge, UK. [7]London School of Hygiene and Tropical Medicine, London, UK. [8]These authors contributed equally: Marcus C. de Goffau, Susanne Lager. [9]These authors jointly supervised this work: D. Stephen Charnock-Jones, Gordon C. S. Smith. *e-mail: jp369@cam.ac.uk; gcss2@cam.ac.uk

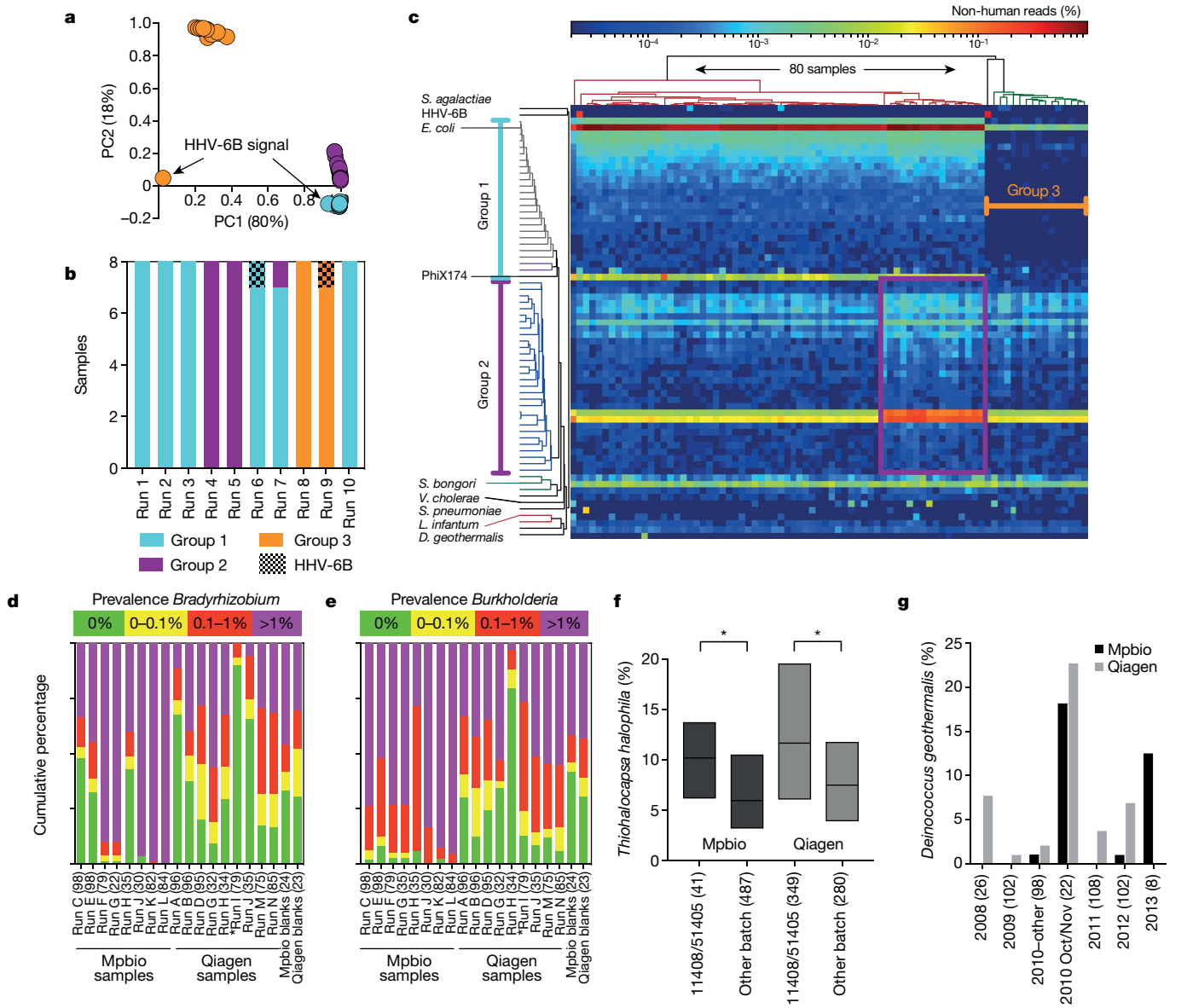


Fig. 1 | Batch effect detection in metagenomic and 16S rRNA amplicon sequencing data, cohort 1 samples. a–c, Summary of metagenomics data. **a**, PCA of summarized genus level identified by Kraken[25] output. **b**, MiSeq sequencing runs ($n = 8$ per run). **c**, Heat map of all non-human read abundance (see Extended Data Fig. 4). **d**, **e**, Read abundance by run and DNA isolation method (Mpbio or Qiagen) in chronological order for *Bradyrhizobium* (**d**) and *Burkholderia* (**e**). Scatterplots are shown in Extended Data Fig. 6. **f**, Associations between *Thiohalocapsa halophila* and Q5 buffer (lot 11508) or Taq polymerase (lot 51405). Interquartile range is shown; centre values denote medians. *$P < 0.001$ (Mann–Whitney *U*-test). **g**, *D. geothermalis* detection (>0.1% reads) by year of delivery. The number of samples in each group in **f** and **g** is shown in parentheses.

bacterial signals was assessed using the kappa statistic, scaled from 0 (no agreement) to 1 (perfect agreement). Only two signals demonstrated agreement (moderate-substantial) between the two methods: *S. agalactiae* and *Deinococcus geothermalis* (Table 1). The results were consistent when using different definitions of positive (Supplementary Table 3) and neither signal was detected in negative controls. The number of positive samples was too small for informative comparison of cases and controls.

Several bacterial signals associated with principal component 2, including the *Caulobacter*, *Methylobacterium* and *Burkholderia* genera, were also detected by 16S rRNA gene sequencing. However, the kappa statistics were low and these signals were also detected in negative controls (Table 1). *Vibrio cholerae* and *Streptococcus pneumoniae* signals were detected using metagenomics in 14 and 11 samples, respectively. However, neither was detected using 16S rRNA sequencing (Table 1). Assembly and analysis of these reads demonstrated that the closest matches were isolates from Bangladesh (PRJEB14661 *V. cholerae*) and the Global Pneumococcal Sequence Project (PRJEB31141 *S. pneumoniae*), which had been sequenced on the same pipelines at the Sanger Institute, indicating that these signals are due to cross-contamination during library preparation or sequencing (the same explanation applies for *Leishmania infantum*, Fig. 1c).

Cohort 2: duplicate 16S rRNA

By combining the data from two independent DNA isolation methods (the MP Biomedical kit, hereafter 'Mpbio', or Qiagen kit), we were able to visualize batch effects using PCA (Extended Data Fig. 5a) or visualize species individually (Fig. 1d–g) and analyse signal reproducibility. For example, *Bradyrhizobium* was detected nearly ubiquitously and in high abundance in some 16S rRNA sequencing runs, but was less frequently detected and in lower abundance in others (Fig. 1d, compare runs K and L with runs I and J). The *Burkholderia* genus, which has been suggested to have a role in PTB[3], had a higher signal in samples isolated using the Mpbio DNA isolation reagents than with the Qiagen kit, and also showed pronounced run-to-run variation (Fig. 1e). Furthermore, both *Bradyrhizobium* and *Burkholderia* were commonly detected in

Table 1 | Comparison of main signals using metagenomics with 16S rRNA amplicon sequencing

Species	Positive signals MG and 16S (79 = max)				MG reads in positive samples	16S reads in positive samples[a]	Concordance MG and 16S kappa score (*P* value)[b]	Part of an MG batch effect[c]	Presence 16S in negative controls Absent/weak/strong ($n = 5$)[d]
	Both[a]	MG only	16S only[a]	Neither					
Salmonella bongori	79	0	0	0	178	54%	NA	No	5/0/0
Escherichia coli	1	78	0	0	18,602	1.2%	0 (–)	Gr. 1 & 2	4/1/0
Shigella (genus)	0	75	0	4	254	NA	0 (–)	Gr. 1 & 2	5/0/0
Salmonella enterica	0	75	0	4	33	NA	0 (–)	Gr. 1 & 2	5/0/0
Cronobacter sakazakii	0	65	0	14	21	NA	0 (–)	Gr. 1 & 2	5/0/0
Bacillus subtilis	0	63	0	16	13	NA	0 (–)	Gr. 1 & 2	5/0/0
Yersinia pseudotuberculosis	0	59	0	20	3	NA	0 (–)	Gr. 1 & 2	5/0/0
Neisseria meningitidis	0	44	0	35	2	NA	0 (–)	Gr. 1 & 2	5/0/0
Bradyrhizobium (genus)	0	79	0	0	125	NA	0 (–)	Gr. 2	5/0/0
Rhodopseudomonas palustris	0	79	0	0	45	NA	0 (–)	Gr. 2	5/0/0
Caulobacter (genus)	12	67	0	0	14	1.4%	0 (–)	Gr. 2	1/3/1
Methylobacterium (genus)	9	69	0	1	8	2.4%	0.003 (0.36)	Gr. 2	1/4/0
Burkholderia (genus)	21	57	0	1	7	1.9%	0.009 (0.27)	Gr. 2	1/4/0
Propionibacterium acnes	66	13	0	0	20	4.8%	0 (–)	No	0/3/2
Streptococcus pneumoniae	0	11	0	68	115	NA	0 (–)	No	5/0/0
Vibrio cholerae	0	14	0	65	46	NA	0 (–)	No	5/0/0
Thiohalocapsa halophila	0	0	71	8	NA	4.2%	0 (–)	No	0/0/5
Stenotrophomonas maltophilia	5	51	1	22	2	1.9%	0.03 (0.24)	No	2/3/0
Acinetobacter baumanii	1	26	0	52	2	2.4%	0.05 (0.08)	No	4/1/0
Micrococcus luteus	1	46	0	32	15	2.0%	0.02 (0.20)	No	4/1/0
Gardnerella vaginalis	0	5	0	74	1	NA	0 (–)	No	4/1/0
Lactobacillus crispatus	0	4	0	75	1	NA	0 (–)	No	5/0/0
Deinococcus geothermalis	1	1	0	77	68	33%	0.66 (<0.0001)	No	5/0/0
Streptococcus agalactiae	3	4	0	72	8	13%	0.58 (<0.0001)	No	5/0/0

The average number of metagenomics (MG) and average percentage of 16S reads in positive samples are shown. Gr., group; NA, not applicable.
[a]16S rRNA amplicon sequencing signals higher than 1% are defined as positive.
[b]One-sided *P* values (kappa statistic).
[c]See Fig. 1 for definition of groups 1 and 2.
[d]Strong signals are defined as more than 1%.

the negative controls. Batch effects based on the use of particular polymerase chain reaction (PCR) reagent lots can also be visualized. For example, the association of *Thiohalocapsa halophila* with either the PCR reagent '5× Q5 buffer' (lot 11408) or 'Q5 Taq polymerase' (lot 51405), both of which were used to process the same 390 samples, is shown in Fig. 1f.

We used the kappa statistic to quantify the level of agreement between 16S rRNA amplicon sequencing of two DNA samples from the mode of delivery. $^*P < 0.05$, $^{***}P < 0.001$, Mann–Whitney *U*-tests were used where values below 1% are regarded as 0%. See Extended Data Fig. 6 for scatterplots. Percentage read count is based on the higher value for given species using Qiagen or Mpbio DNA isolation kit (using all 498 samples).

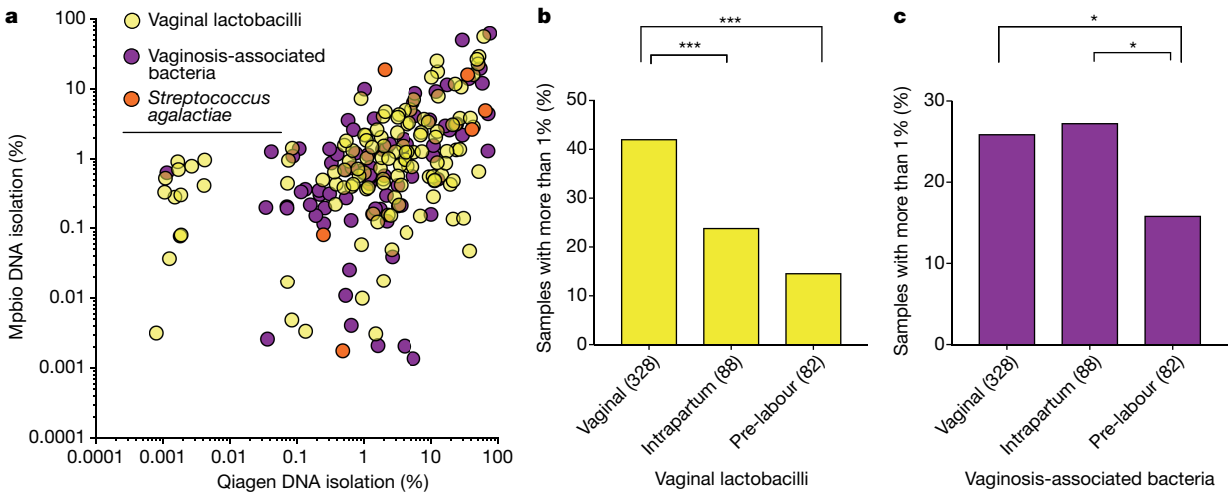


Fig. 2 | Mode of delivery and detection of vaginal bacteria by 16S rRNA amplicon sequencing. a, Concordant detection of vaginal lactobacilli and a combination of all vaginosis-associated bacteria by both Qiagen (*x* axis) and Mpbio (*y* axis) results in Spearman's rho correlation coefficients of 0.37 and 0.59, respectively, when analysing the top right quadrant only (>0.1%). **b**, **c**, Comparisons between vaginally associated bacteria and

Table 2 | Simplified overview on the nature of bacterial findings

	Signals					
	Independent of:					
	DNA extraction batch[a]	Date of delivery[b]	Mode of delivery	Not in negative controls[c]	Sample-associated[d]	Verified by meta-genomics[e]
Capable pathogens						
Streptococcus agalactiae	✓	✓	✓	✓	✓	✓
Listeria monocytogenes[f]	✓	✓	✓	✓	✓	–
Vaginal lactobacilli						
Lactobacillus crispatus	✓	✓	–	~	✓	~
Lactobacillus iners	✓	✓	–	~	✓	–
Lactobacillus gasseri	✓	✓	–	✓	✓	–
Lactobacillus jensenii	✓	✓	–	~	✓	–
Vaginosis-associated bacteria						
Gardnerella vaginalis	✓	✓	–	~	✓	–
Atopobium vaginae	✓	✓	–	~	✓	–
Ureaplasma (genus)	✓	✓	–	✓	✓	–
Prevotella bivia	✓	✓	–	~	✓	–
Prevotella amnii	✓	✓	–	✓	✓	–
Prevotella timonensis	✓	✓	–	~	✓	–
Aerococcus christensenii	✓	✓	–	✓	✓	–
Streptococcus anginosus	✓	✓	–	~	✓	–
Sneathia sanguinegens	✓	✓	–	✓	✓	–
Megasphaera elsdenii	✓	✓	–	~	✓	–
Faecal-associated bacteria						
Bacteroides (genus)	✓	✓	–	~	✓	–
Faecalibacterium prausnitzii	✓	✓	–	~	✓	–
Roseburia faeces	–	✓	–	~	✓&–	–
Coriobacterium sp.	✓	✓	–	~	✓	–
Collinsella intestinalis	✓	✓	–	+	✓	–
Suspected oral origin						
Fusobacterium nucleatum	✓	✓	✓	~	✓	–
Streptococcus mitis	✓	✓	✓	~	✓	–
Streptococcus vestibularis	–	✓	✓	~	✓&–	–
Genuine reagent contaminants						
Acinetobacter baumanii[f]	–	✓	✓	~	–	~
Thiohalocapsa halophila	–	✓	✓	–	–	–
Propionibacterium acnes	–	✓	✓	–	–	–
Stenotrophomonas maltophilia	–	✓	✓	–	–	–
Bradyrhizobium japonicum	–	✓	✓	–	–	–
Melioribacter roseus	–	✓	✓	–	–	–
Pelomonas (genus)	–	✓	✓	–	–	–
Methylobacterium (genus)	–	✓	✓	–	–	–
Aquabacterium (genus)	–	✓	✓	–	–	–
Sediminibacterium (genus)	–	✓	✓	–	–	–
Desulfovibrio alkalitolerans	–	✓	✓	–	–	–
Delftia tsuruhatensis	–	✓	✓	–	–	–
Streptococcus pyogenes	–	✓	✓	~	–	–
Burkholderia multivorans	–	✓	✓	–	–	–
Caulobacter (genus)	–	✓	✓	–	–	–
Steroidobacter sp. JC2953	–	✓	✓	–	–	–
Afipia (genus)	–	✓	✓	–	–	–
Burkholderia silvatlantica	–	✓	✓	–	–	–
Lysinimicrobium mangrove	–	✓	✓	–	–	–
Bradyrhizobium elkanii	–	✓	✓	–	–	–
Achromobacter xylosoxidans	–	✓	✓	–	–	–
Corynebacterium tuberculostearicum	–	✓	✓	–	–	–
Rhodococcus fascians	✓	–	✓	~	✓	–
Sphingobium rhizovicinum	✓	–	✓	~	✓	–
Methylobacterium organophilum	✓	–	✓	~	✓	–
Deinococcus geothermalis[f]	✓	–	✓	✓	✓	✓

[a]Includes batch effects caused by different DNA isolation kits, PCR reagents and MiSeq run.
[b]See Figs. 1g, 2d for details.
[c]A tick '✓' indicates absence; '~' indicates detection (any percentage) in less than 20% of negative controls.
[d]Detection of signal in corresponding Qiagen and Mpbio DNA isolations. '✓&–' indicates that signals from these operational taxonomic units are sample-associated in most 16S runs, but reagent contaminants in others. See Supplementary Table 4 for details.
[e]See Table 1 and Supplementary Table 3. A '~' indicates some level of concordance was detected using a different 16S threshold.
[f]The presence or absence of verification should be interpreted with caution, as indicated by examples.

the same patient extracted using the two different kits (Supplementary Table 4). The majority of the most-prevalent bacterial groups had low kappa scores and there was a low correlation between the magnitude of the signals comparing the two DNA extraction methods (Extended Data Fig. 5b). Moreover, these signals also demonstrated notable batch effects using PCA (Extended Data Fig. 5a). Interestingly, four ecologically unexpected bacterial groups of high prevalence exhibited a fair level of concordance (*Rhodococcus fascians*, *Sphingobium rhizovicinum*, *Methylobacterium organophilum* and *D. geothermalis*). Further analysis demonstrated a temporal pattern of these signals (Fig. 1g). All placental samples were washed in sterile PBS to remove surface contamination, such as maternal blood, and the temporal pattern of these bacterial signals is consistent with them being derived from batches of this reagent. Some ecologically plausible species, such as *S. agalactiae* and *Listeria monocytogenes*, vaginal lactobacilli, vaginosis-associated bacteria, faecal bacteria and some bacteria of probable oral origin had modest to high kappa scores, indicating that they were sample-associated signals. In contrast to the laboratory contaminants, the signals for these bacterial groups correlated when comparing the two DNA extraction methods (Fig. 2a) and were not associated with batch effects identifiable using PCA. Sample-associated signals (non-reagent contaminants) of a few species not typically associated with a vaginal or rectal habitat but with the oral habitat were detected, such as *Streptococcus mitis*, *Streptococcus vestibularis* and *Fusobacterium nucleatum*. However, it was only a very small minority of samples that exhibited these signals (below that of *S. agalactiae*) and none of these oral signals was identified by metagenomic analysis of pre-labour Caesarean section samples (cohort 1).

Delivery-associated signals

Vaginal organisms (lactobacilli and vaginosis-associated bacteria) were more abundant than *S. agalactiae* in cohort 2 (vaginal, intrapartum and pre-labour Caesarean section deliveries) but less abundant than *S. agalactiae* in cohort 1 (pre-labour Caesarean section deliveries only). Hence, we next examined the relationship between the mode of delivery and the 16S rRNA signal. Vaginal lactobacilli (*Lactobacillus iners*, *Lactobacillus crispatus*, *Lactobacillus gasseri* and *Lactobacillus jensenii*) were found more frequently and in higher numbers in vaginally delivered placentas than in placentas delivered via intrapartum or pre-labour Caesarean section (Fig. 2b), irrespective of the DNA isolation method (Extended Data Fig. 7a, b). Vaginosis-associated bacteria were found at approximately the same frequency in vaginal and intrapartum Caesarean section samples, but significantly less frequently in pre-labour Caesarean section samples (Fig. 2c). A heat map generated using the Spearman rho correlation coefficients of all abundant and relevant bacterial groups generated a cluster of vaginally associated bacteria, representative of vaginal community group IV[16], which reflects sample contamination during labour and delivery (Extended Data Fig. 8). The other clusters represented the contamination signatures of the two different DNA extraction kits and a fourth cluster reflected contamination associated with the date of collection of the placental biopsies (2012–2013).

Genuine signals and pregnancy outcome

The presence of *S. agalactiae* was analysed with respect to clinical outcome (SGA, pre-eclampsia, PTB) as it was the only organism that met all of the criteria of a genuine placenta-associated bacterial signal (Table 2). There was no association with SGA, pre-eclampsia or PTB (Fig. 3). Exploratory analysis of the 16S amplicon sequencing data of all sample-associated signals, including delivery-associated bacteria, showed that *S. mitis* and *F. nucleatum* were not associated with adverse pregnancy outcome (Supplementary Table 5). Of note, however, were the significant associations of the delivery-associated bacteria *L. iners* with pre-eclampsia and *Streptococcus anginosus* and the *Ureaplasma* genus with PTB (Fig. 3, Supplementary Table 5 and Extended Data Fig. 9). In one placental sample from a preterm birth, a strong

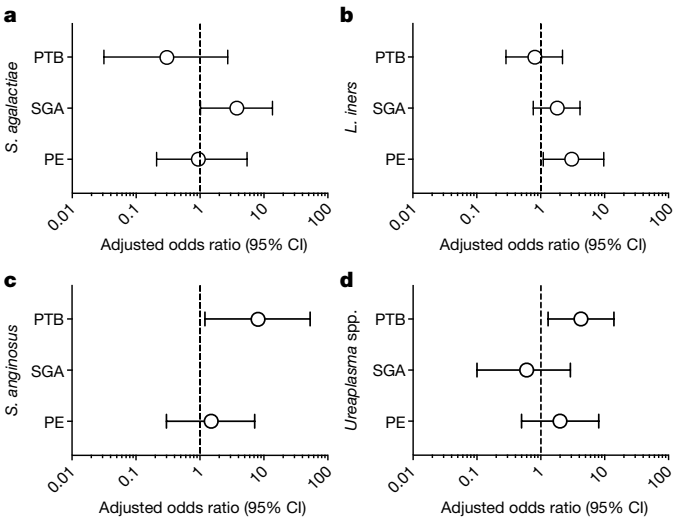


Fig. 3 | Bacterial signals and adverse pregnancy outcome. a–d, Adjusted odds ratios for the association of *S. agalactiae* (**a**), *L. iners* (**b**), *S. anginosus* (**c**) and *Ureaplasma* spp. (**d**) with PTB, SGA and pre-eclampsia (PE). Pre-eclampsia and SGA both had 100 matched cases and controls. The PTB analysis included 56 preterm cases and 136 unmatched controls (all vaginally delivered). Odds ratios were adjusted for clinical characteristics by logistic regression. The odds ratio and its confidence interval (CI) cannot be calculated for *S. anginosus* and SGA because one of the discordant values is zero. See Supplementary Table 5 for further details.

L. monocytogenes signal was found (7% and 52% of all reads with Mpbio and Qiagen, respectively).

Validating *Streptococcus agalactiae*

A nested PCR and quantitative PCR (qPCR) approach targeted towards the *sip* gene, which encodes the surface immunogenic protein (SIP) of *S. agalactiae*, was used to verify its presence in 276 placental samples for which a 16S sequencing result was available. In total, 7 out of 276 samples were positive using PCR–qPCR and all seven were also positive (more than 1%) by 16S analysis. A total of 14 samples were positive by 16S sequencing but not by PCR–qPCR, no sample was positive using PCR–qPCR and negative by 16S, and 255 samples were negative by both methods. This yielded a kappa statistic of 0.48, indicating moderate agreement and a P value of 9.7×10^{-21}. We conclude that the detection of *S. agalactiae* by 16S rRNA amplification was verified by two further independent methods (metagenomics and PCR–qPCR) and the level of agreement in both cases was well above what could be expected by chance. It remains to be determined why some samples were positive for *S. agalactiae* by 16S sequencing but negative by the PCR–qPCR method. Generally, the latter would be considered more sensitive, particularly in samples with a higher microbial biomass, owing to the complex amplification kinetics when a large number of diverse 16S template molecules are present. However, in the absence of other bacterial signals, it is possible that 16S sequencing is more sensitive for detecting very small numbers of *S. agalactiae*, as the genome of the organism has seven copies of the 16S rRNA gene, but only one copy of *sip*[17].

Discussion

We studied placental biopsies from a total of 537 women, including 318 cases of adverse pregnancy outcome and 219 controls, using multiple methods of DNA extraction and detection, and drew several important conclusions. First, we found that the biomass of bacterial sequences in DNA extracted from human placenta was extremely small. Second, the major source of bacterial DNA in the samples studied was contamination from laboratory reagents and equipment. Third, both metagenomics and 16S amplicon sequencing were capable of detecting a very low amount of a spiked-in signal. Fourth, samples of placental tissue

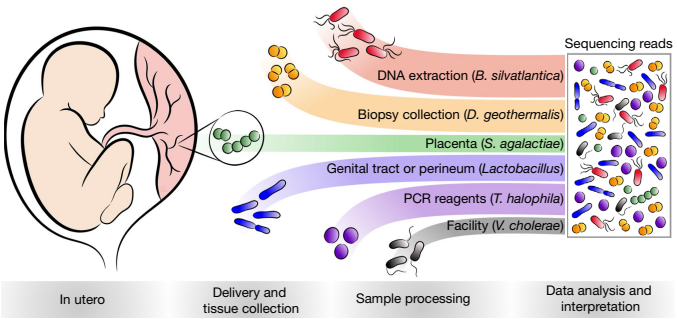


Fig. 4 | Sources of bacterial signals detected in human placental samples. Bacteria may sometimes be present in utero, such as *S. agalactiae*. Bacteria or bacterial DNA also frequently contaminate the placenta during labour and delivery (for example, *Lactobacillus*), during sample collection (for example, *D. geothermalis*), and during sample processing (for example, *B. silvatlantica* and *T. halophila*). Contamination may also occur during library preparation or sequencing from other projects carried out at the facility (for example, *V. cholerae* in the metagenomic sequencing).

become contaminated during the process of labour and delivery, even when they were dissected from within the placenta. Finally, the only organism for which there was strong evidence that it was present in the placenta before the onset of labour was *S. agalactiae*. It was not part of any batch effect, it was detected by three methods, there was a statistically significant level of agreement between 16S amplicon sequencing and both metagenomics ($P = 1.5 \times 10^{-8}$) and a targeted PCR–qPCR assay ($P = 9.7 \times 10^{-21}$), none of 47 negative controls analysed by 16S sequencing was positive for *S. agalactiae*, and there was no association with mode of delivery (Extended Data Fig. 7). However, there was no significant association between the presence of the organism and pre-eclampsia, SGA or PTB. Exploratory analysis of other signals did demonstrate an association between PTB and the presence of *Ureaplasma* reads (>1%), consistent with previous studies[18], but this was probably the result of ascending uterine infection. We conclude that bacterial placental infection is not a major cause of placentally related complications of human pregnancy and that the human placenta does not have a resident microbiome.

The finding of *S. agalactiae* in the placenta before labour could be of considerable clinical importance. Perinatal transmission of *S. agalactiae* from the mother's genital tract can lead to fatal sepsis in the infant. It is estimated that routine screening of all pregnant women for the presence of *S. agalactiae* and targeted use of antibiotics prevents 200 neonatal deaths per year in the United States[19]. Our findings identify an alternative route for perinatal acquisition of *S. agalactiae*. Further studies will be required to determine the association between the presence of the organism in the placenta and fetal or neonatal disease. However, if such a link was identified, rapid testing of the placenta for the presence of *S. agalactiae* might allow targeting of neonatal investigation and treatment. Our work also sheds light on the possible routes of fetal colonization. Although we see no evidence of a placental microbiome, the frequency of detection of vaginal bacteria in the placenta increased after intrapartum Caesarean section, suggesting ascending or haematogenous spread. Similarly, haematogenous spread as the result of transient bacteraemia could potentially explain the presence of the small number of sample-associated oral bacterial signals[14]. Such spread could lead to fetal colonization immediately before delivery.

We identified five different patterns of contamination (Fig. 4)—namely, contamination of the placenta with real bacteria during the process of labour and delivery (Fig. 2); contamination of the biopsy when it was washed with PBS; contamination of DNA during the extraction process; contamination of reagents used to amplify the DNA before sequencing; and contamination from the reagents or equipment used for sequencing. Using 16S rRNA amplicon sequencing, the positive control (*S. bongori*) accounted for more than half of the reads, indicating that the method is highly sensitive. However, when the method is applied to samples with little or no biomass, these sources of contamination can lead to apparent signals, hence it is crucial to use a method that allows differentiation between true bacterial signals and these sources of contamination (see Supplementary Information 1 for further technical discussion).

In conclusion, in a study of 537 placentas carefully collected, processed and analysed to detect real bacterial signals, we found no evidence to support the existence of a placental microbiome and no significant relationship between placental infection with bacteria and the risk of pre-eclampsia, SGA and preterm birth. However, we identified an important pathogen, *S. agalactiae*, in the placenta of approximately 5% of women before the onset of labour.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1451-5.

Received: 16 January 2019; Accepted: 28 June 2019;
Published online 31 July 2019.

1. Brosens, I., Pijnenborg, R., Vercruysse, L. & Romero, R. The "Great Obstetrical Syndromes" are associated with disorders of deep placentation. *Am. J. Obstet. Gynecol.* **204**, 193–201 (2011).
2. Aagaard, K. et al. The placenta harbors a unique microbiome. *Sci. Transl. Med.* **6**, 237ra65 (2014).
3. Antony, K. M. et al. The preterm placental microbiome varies in association with excess maternal gestational weight gain. *Am. J. Obstet. Gynecol.* **212**, 653.e1–653.e16 (2015).
4. Collado, M. C., Rautava, S., Aakko, J., Isolauri, E. & Salminen, S. Human gut colonisation may be initiated in utero by distinct microbial communities in the placenta and amniotic fluid. *Sci. Rep.* **6**, 23129 (2016).
5. Perez-Muñoz, M. E., Arrieta, M. C., Ramer-Tait, A. E. & Walter, J. A critical assessment of the "sterile womb" and "in utero colonization" hypotheses: implications for research on the pioneer infant microbiome. *Microbiome* **5**, 48 (2017).
6. Salter, S. J. et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* **12**, 87 (2014).
7. Jervis-Bardy, J. et al. Deriving accurate microbiota profiles from human samples with low bacterial content through post-sequencing processing of Illumina MiSeq data. *Microbiome* **3**, 19 (2015).
8. de Goffau, M. C. et al. Recognizing the reagent microbiome. *Nat. Microbiol.* **3**, 851–853 (2018).
9. Lauder, A. P. et al. Comparison of placenta samples with contamination controls does not provide evidence for a distinct placenta microbiota. *Microbiome* **4**, 29 (2016).
10. Leiby, J. S. et al. Lack of detection of a human placenta microbiome in samples from preterm and term deliveries. *Microbiome* **6**, 196 (2018).
11. Theis, K. R. et al. Does the human placenta delivered at term have a microbiota? Results of cultivation, quantitative real-time PCR, 16S rRNA gene sequencing, and metagenomics. *Am. J. Obstet. Gynecol.* **220**, 267.e1–267.e39 (2019).
12. Leon, L. J. et al. Enrichment of clinically relevant organisms in spontaneous preterm delivered placenta and reagent contamination across all clinical groups in a large UK pregnancy cohort. *Appl. Environ. Microbiol.* **84**, e00483-e18 (2018).
13. Sovio, U., White, I. R., Dacey, A., Pasupathy, D. & Smith, G. C. S. Screening for fetal growth restriction with universal third trimester ultrasonography in nulliparous women in the Pregnancy Outcome Prediction (POP) study: a prospective cohort study. *Lancet* **386**, 2089–2097 (2015).
14. Hornef, M. & Penders, J. Does a prenatal bacterial microbiota exist? *Mucosal Immunol.* **10**, 598–601 (2017).
15. Leong, H. N. et al. The prevalence of chromosomally integrated human herpesvirus 6 genomes in the blood of UK blood donors. *J. Med. Virol.* **79**, 45–51 (2007).
16. Ravel, J. et al. Vaginal microbiome of reproductive-age women. *Proc. Natl Acad. Sci. USA* **108** (suppl. 1), 4680–4687 (2011).
17. Glaser, P. et al. Genome sequence of *Streptococcus agalactiae*, a pathogen causing invasive neonatal disease. *Mol. Microbiol.* **46**, 1499–1513 (2002).
18. Abele-Horn, M., Scholz, M., Wolff, C. & Kolben, M. High-density vaginal *Ureaplasma urealyticum* colonization as a risk factor for chorioamnionitis and preterm delivery. *Acta Obstet. Gynecol. Scand.* **79**, 973–978 (2000).
19. Schrag, S. J. et al. Group B streptococcal disease in the era of intrapartum antibiotic prophylaxis. *N. Engl. J. Med.* **342**, 15–20 (2000).

METHODS

Ethics. This study is in compliance with all relevant ethical regulations. The Pregnancy Outcome Prediction study (POPs) was approved by the Cambridgeshire 2 Research Ethics Committee (reference number 07/H0308/163). The study and the characteristics of the eligible and participating women have been previously described in detail[13,20]. In brief, 4,212 nulliparous women with a singleton pregnancy were followed through from their first ultrasound scan to delivery. At the time of delivery, placental samples were obtained using a standardized protocol by a team of trained technicians, in which most samples were obtained within 3 h of delivery (interquartile range: 0.3–8.4 h). All participants gave written informed consent for the study and for subsequent analysis of their samples.

Patient selection. For cohort 1, cases of SGA (≤fifth percentile based on customized birth weight[21]; $n = 20$) or pre-eclampsia (according to the 2013 ACOG (The American College of Obstetricians and Gynaecologists) Guidelines[22]; $n = 20$) were matched one-to-one with healthy controls ($n = 40$). Only deliveries by pre-labour Caesarean section were included in this cohort. The cases and controls were matched as closely as possible for maternal body mass index, maternal age, gestational age, sample collection time, maternal smoking, and fetal sex. Clinical characteristics are presented in Supplementary Table 1.

For cohort 2, cases of SGA (≤fifth customized birth weight percentile[21]; $n = 100$) or pre-eclampsia (2013 ACOG guidelines[22]; $n = 100$) were selected. The cases were matched one-to-one with healthy controls ($n = 198$, two controls were used twice). All deliveries were at term (≥37 weeks gestation). The same matching criteria as in the first cohort were used with the addition of an absolute match for mode of delivery. Placentas from 100 preterm births (<37 weeks gestation) deliveries were also included in the study (clinical characteristics in Supplementary Table 2). Flow charts describing the two cohorts and subsequent sample-processing and analysis steps are presented in Extended Data Fig. 1.

Placenta collection. Placentas were collected after delivery and the procedure has previously been described in detail[20]. We confined our sampling to the placental terminal villi (fetal tissue). We chose this as the villi are the site of exchange, across the vasculosyncytial membrane, between the fetus and mother. This location is the closest interface between the fetus with the mother's blood and tissues. If the placenta was colonized, one would expect bacteria to ascend the genital tract (local infiltration) or to come from the mother's blood (haematogenous). Hence, we believe that this would be the most plausible site for bacteria to be found. Villous tissue was obtained from four separate lobules of the placenta after trimming to remove adhering decidua from the basal plate. The tissue in the selected areas had no visible damage, haematomas, or infarctions. To remove maternal blood, the selected tissue samples were rinsed in chilled sterile PBS (Oxoid Phosphate Buffered Saline Tablets, Dulbecco A; Thermo Fisher Scientific) dissolved in ultrapure water (ELGA Purelab Classic 18MΩ.cm). After initial collection, all placental samples were frozen in liquid nitrogen and stored at −80 °C until further processing. For DNA isolation, approximately 25 mg of villous tissue (combined weight obtained from fragments of all four biopsy collection points) was cut from the stored tissue. To reduce the risk of environmental contamination of the samples, the entire experimental procedure was carried out in a class 2 biological safety cabinet (tissue cutting, DNA isolation, setting up PCR reactions). The tissue was cut with single-use sterile forceps and scalpel. Each matched case–control pair was processed in parallel on the same day for each step of the entire experimental procedure (tissue cutting, DNA isolation, setting up PCR reactions). Also, the same lot of laboratory reagents was used for each pair. For each lot of laboratory reagents, negative controls were included (described in detail below).

DNA isolation from cohort 1. DNA was isolated from placental tissue with the Qiagen Qiaamp DNA mini kit (51304; Qiagen) according to the manufacturer's instructions with the addition of a freeze–thaw cycle after the overnight tissue lysis. Before DNA isolation, intact *S. bongori* was added to the placental tissue (1,100 CFUs, described in detail below). The placental tissue with added *S. bongori* was lysed in a proteinase-K-based solution (100 µl buffer ATL (Qiagen), 80 µl of *S. bongori*, 20 µl proteinase K) overnight (18 h at 56 °C) and thereafter freeze–thawed once. After the thawed samples were brought to room temperature, RNA was removed with the addition of 4 µl RNase A (Qiagen, 19101) and incubated at room temperature for 2 min. Spin-filtering and washing of the DNA was carried out according to the manufacturer's instructions. The DNA was eluted from the spin column with 200 µl buffer AE (Qiagen) after a 5 min incubation (the elution step was repeated once with another 200 µl buffer AE and 5 min incubation). To prevent accidental cross-contamination between samples, gloves were changed between handling each sample. Throughout the protocol (DNA extraction, primer aliquoting, 16S rRNA gene amplification and library preparation), nuclease-free plastics were used (unless supplied with kit): PCR clean 2.0 and 1.5 ml DNA LoBind Tubes (Eppendorf), and nuclease-free filter tips (TipONE sterile filter tips, STARLAB). For each box of DNA isolation kit used, extraction blanks were carried out. These DNA extraction blanks, or negative controls, contained only the reagents from each DNA isolation kit (no added biological material) and were subjected to the complete DNA extraction procedure: tissue homogenization, matrix binding, spin-filtering, washing, and elution of nucleic acids. The negative controls were subjected to the entire analysis protocol alongside the placental samples: DNA isolation, 16S rRNA gene PCR amplification, sequencing and data analysis.

Positive control. As a positive control, a known amount of intact *S. bongori* (strain NCTC-12419) was added to each of the placental tissue samples in cohort 1 before DNA isolation ($n = 80$). *S. bongori* was incubated with shaking overnight at 37 °C in LB broth. When the OD_{600} reached 0.9 (approximately equivalent to 7.2×10^8 bacteria per ml, measured with a Ultrospec 10 Cell Density Meter, GE Healthcare) the culture was chilled on ice. To minimize bacterial growth outside of the shaking incubator, all cultures and dilutions were kept on ice. To increase the proportion of live bacteria added as positive controls, 1 ml of the *S. bongori* suspension was diluted in 14 ml fresh LB broth (OD_{600} was 0.06) and incubated with shaking (1.5 h at 37 °C; OD_{600} was 0.8). The *S. bongori* culture was then serially diluted to an estimated concentration of 1,000 *S. bongori* per 80 µl, which was used to spike the placental samples. To determine the actual number of CFUs added to the placental samples, the *S. bongori* suspension was further diluted and aliquots cultured on LB plates overnight (37 °C). The number of colonies was counted. On the basis of three plates with distinct individual colonies (between 29 and 205 colonies per plate), the number of *S. bongori* added to each placental tissue sample was calculated to be 1,100 CFUs.

DNA isolation from cohort 2. DNA was isolated twice from each placenta using two different extraction kits. The DNA isolations were carried out in accordance with respective manufacturer's instructions, with the addition of two extra washes in the MP Biomedical kit.

For the Qiagen Qiaamp DNA mini kit (Qiagen, 51304), the placental tissue was digested in a proteinase-K-based solution (100 µl buffer ATL, 80 µl PBS, 20 µl proteinase K) for at least 3 h. Then, 4 µl of RNase A (Qiagen, 19101) was added to the tissue lysate and incubated at room temperature for 2 min. Spin-filtering and washing of the DNA was carried out according to the manufacturer's instructions. The DNA was eluted from the spin column with 200 µl buffer AE after a 5 min incubation (the elution step was repeated once with another 200 µl buffer AE and 5 min incubation).

For the MP Biomedical Fast DNA Spin kit (MP Biomedical, 116540600), the placental tissue was homogenized in 1.0 ml of CLS-TC solution by bead-beating (Lysing Matrix A tubes, 40 s, speed 6.0 on a FastPrep-24, MP Biomedical). After spinning the samples, equal volumes of the supernatant were combined with Binding Matrix. The mixture was transferred to a spin filter, after spin filtering the DNA was washed three times with SEWS-M. The DNA was eluted by re-suspending the Binding Matrix in 100 µl DES buffer, incubating the tubes at 55 °C for 5 min before recovering the DNA by centrifugation.

The same measures to prevent contamination of the samples as described in the cohort 1 DNA isolation section were taken. Extraction blanks were generated for each box/lot of both DNA isolation kits in a similar manner as was done for cohort 1. DNA concentrations were determined by Nanodrop Lite (Thermo Fisher Scientific).

Metagenomic sequencing. Sample processing for the metagenomics analysis was performed exactly as previously described[23]. In brief, the NEB Ultra II custom kit (New England Biolabs) was used for library generation, and samples were then sequenced on the Illumina HiSeq X Ten platform (150 base pairs, paired end) in 10 runs (flowcells) of 8 samples (lanes) each. The sequencing coverage was designed to generate more than 30-fold coverage of the human chromosomal DNA in each sample.

16S rRNA gene amplification. For detection of the bacterial 16S rRNA gene, PCR amplification of the V1–V2 region was performed using V1 primers with four degenerate positions to optimize coverage as previously recommended[24]. The V1–V2 amplicon is relatively short (~260 bp) and, with paired-end reads, almost all of the amplified product is sequenced on both strands and thus at higher accuracy. This is not the case with the longer V1–V3 amplicon. This region has also been used in other studies of the placental microbiome[10]. The following barcoded primers were used forward-27: 5′-AATGATACGGCGACCACCGAGATCTACAC*nnnnnnnnnnnn*ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNAGMGTTYGATYMTGGCTCAG-3′ and reverse-338: 5′-CAAGCAGAAGACGGCATACGAGAT *nnnnnnnnnnnn* GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNG CTGCCTCCCGTAGGAGT. The *n*-string represents unique 12-mer barcodes used for each sample studied and distinct indexes were used at both the 5′ and 3′ ends of the amplicons. The primers were purchased from Eurofins Genomics. Before aliquoting, the cabinet and pipettes were cleaned with DNA AWAY Surface Decontaminant. The primers were diluted in Tris-EDTA buffer (Sigma-Aldrich) in PCR clean nuclease-free DNA LoBind Tubes (Eppendorf) with nuclease-free filter tips (TipONE sterile filter tips, STARLAB). The PCR amplification was carried out in quadruplicate reactions for each sample on a SureCycler 8800 Thermal Cycler (Agilent Technologies) with high-fidelity Q5 polymerase (M0491L; New

England Biolabs), dNTP solution mix (N0447L, New England Biolabs), and UltraPure DNase/RNase-Free Water (Thermo Fisher Scientific) in 0.2 ml PCR strips (STARLAB). Amplification was performed with 500 ng DNA per reaction, and the final primer concentration was 0.5 μM. The PCR amplification profile was an initial step of 98 °C for 2 min followed by 10 cycles of touch-down (68 to 59 °C; 30 s), and 72 °C (90 s), followed by 30 cycles of 98 °C (30 s), 59 °C (30 s), and 72 °C (90 s). After completion of cycling, the reactions were incubated for 5 min at 72 °C. After completion of the PCR, the four replicates of each sample were pooled, cleaned up with AMPure XP beads (A63881; Beckman Coulter) and eluted in Tris-EDTA buffer (Sigma-Aldrich). DNA concentration was determined by Qubit Fluorometric Quantitation (Q32854; Invitrogen). Equimolar pools of the PCR amplicons were run on 1% agarose/TBE gels and ethidium bromide used to visualize the DNA. The DNA bands were excised and cleaned up with a Wizard SV Gel and PCR Clean-Up System (Promega UK). The equimolar pools were sequenced on the Illumina MiSeq platform using paired-end 250 cycle MiSeq Reagent Kit V2 (Illumina).

Bioinformatic analysis of metagenomics data. Bioinformatic analysis first required removal of human reads followed by identification of the species of non-human reads. KneadData (http://huttenhower.sph.harvard.edu/kneaddata) is a tool designed to perform quality control on metagenomic sequencing data, especially data from microbiome experiments, and we used this to remove the human reads. Forward and reverse reads from each sample were filtered using KneadData (v.0.6.1) with the following trimmomatic options: HEADCROP9, SLIDINGWINDOW:4:20, MINLEN: 100. A custom Kraken[25] reference database (v.0.10.6) was built, using metagm_build_kraken_db and -max_db_size 30, to detect any bacterial, viral and potential non-human eukaryotic signals. This custom Kraken reference database included both the default bacterial and viral libraries, and an accessions.txt file was supplied (via -ids_file) containing a diverse array of organisms chosen from all sequenced forms of eukaryotic life (see Supplementary Table 6 for accession numbers). This wide array was chosen to both detect potentially relevant unknown organisms, but also to identify additional human reads that had not been mapped to the human reference genome. In the metagenomic data, various non-human eukaryotic signals were identified by Kraken in every placental sample at a similar percentage, and were mostly assigned to *Pan paniscus* (Supplementary Table 6). As a verification, reads mapping to eukaryotic species were extracted (Supplementary Information 1) and contigs were assembled. These were analysed using BLASTN and were indeed identified as human. This indicates that these (often lower quality or repetitive) eukaryotic reads are in fact human reads that were not removed by mapping against the human reference genome. An exception to this was that in 17 samples an elevated number of reads were assigned to *Danio rerio* (zebrafish) and *Sarcophilus harrisii* (Tasmanian devil), both of which had been sequenced on the Sanger Institute pipeline. Kraken was run using the metagm_run_kraken option. All human-derived signals (eukaryotic non-fungal reads found in every placental sample at a similar percentage) were removed before further analysis. See Source Data of Fig. 1a–c for abundance information. The origins of *Streptococcus pneumonia* and *Vibrio cholerae* reads were analysed by extracting their respective reads as identified by the Kraken using custom scripts (Supplementary Information 1), performing an assembly on these reads using Spades (v.3.11.0)[26] and by using BLAST (blastn, database: others)[27] to find the closest match. The first step of the strain level analysis of *E. coli* reads to find the closest *E. coli* reference genome match was identical to the steps described above. Subsequently, *E. coli* reads were mapped against *E. coli* WG5 (GenBank: CP02409.1) using BWA (v.0.7.17-r1188)[28] and visualized using Artemis (v.16.0.0)[29]. *E. coli* reads were both analysed per sample and by combining all *E. coli* reads from all samples together.

Bioinformatic analysis of 16S rRNA gene amplicon data. To analyse all 14 16S rRNA amplicon data together using the MOTHUR (v.1.40.5) MiSeq SOP[30] and the Oligotyping (v.2.1) pipeline[31], the data from each individual run were initially individually processed in the MOTHUR pipeline as described below. All of the reads need to be aligned together as a requirement of the Oligotyping pipeline so after the most memory intensive-filtering steps had been performed, they were combined and processed again. Modifications to the MOTHUR MiSeq SOP are as follows: the 'make.contigs' command was used with no extra parameters on each individual run. The assembled contigs were taken out from the MOTHUR pipeline and the four poly NNNNs present in the adaptor/primer sequences were removed using the '-trim_left 4' and '-trim_right 4' parameters in the PRINSEQ-lite (v.0.20.3) program[32]. The PRINSEQ trimmed sequences were used for the first 'screen.seqs' command to remove ambiguous sequences and sequences containing homopolymers longer than 6 bp. In addition, any sequences longer than 450 bp or shorter than 200 bp were removed. Unique reads ('unique.seqs') were aligned ('align.seqs') using the Silva bacterial database 'silva.nr_v123.align'[33] with flip parameter set to true. Any sequences outside the expected alignment coordinates ('start=1046', 'end=6421') were removed. The correctly aligned sequences were subsequently filtered ('filter.seqs') with 'vertical=T' and 'trump='. The filtered sequences were de-noised by allowing three mismatches in the "pre.clustering" step and chimaeras were removed using Uchime with the dereplicate option set to 'true'. The chimaera-free sequences were classified using the Silva reference database 'silva.nr_v123.align' and the Silva taxonomy database 'silva.nr_v123.tax' and a cut-off value of 80%. Chloroplast, mitochondria, unknown, archaea, and eukaryota sequences were removed. All reads from each sample were subsequently renamed, placing the sample name of each read in front of the read name. The 'deunique.seqs' command, which creates a redundant fasta file from a fasta and name file, was performed before concatenating all of the data of all 14 16S runs together using the 'merge.files' command, which was done on both the fasta and the group files. The 'unique.seqs' command was again used before again aligning all reads as described previously before finishing the MOTHUR pipeline with the 'deunique.seqs' command.

Oligotyping and species identification. After the MOTHUR pipeline, the redundant fasta file, which now only contains high-quality aligned fasta reads, was subsequently used for oligotyping using the unsupervised minimum entropy decomposition (MED) for sensitive partitioning of high-throughput marker gene sequences[31]. A minimum substantive abundance of an oligotype (-M) was defined at 1,000 reads and a maximum variation allowed (-V) was set at 3 using the command line 'decompose 14runs.fasta -M 1000 -V 3 -g –t'. The node representative sequence of each oligotype (OTP) was used for species profiling using the ARB program (v.5.5-org-9167)[34]. For ARB analysis, we used a customized version of the SILVA SSU Ref database (NR99, release 123) that was generated by removing uncultured taxa. Oligotype abundances are provided in Supplementary Information 2 and additional metadata, for example, contamination identification via PCA (Extended Data Fig. 3), is provided in the Source Data.

Sensitivity analysis. To compare 16S rRNA amplicon sequencing and metagenomics sensitivity, the *S. bongori* signals (positive control) spiked into cohort 1 were analysed (Extended Data Fig. 2a, b). In 16S rRNA amplicon sequencing analysis 1,100 CFUs of *S. bongori* resulted in an average of 33,000 *S. bongori* reads (~54%). Thus, the remaining bacterial signal (reagent contamination background plus other signals) contributes the remaining 46% of the reads. This is approximately equivalent to another 937 *S. bongori* CFUs (1,100/(54/46)). Thus, if there are 937 bacteria in the sample (everything except the spike), this should produce a signal of 100% when there are no spiked-in bacteria present. Thus, the sensitivity of this assay in cohort 2, which did not contain a spike, is 0.106% of sequencing reads per CFUs (100%/937 CFUs). However, although an average of 54% *S. bongori* reads were detected in all spiked samples, it can be reasoned that samples with the highest *S. bongori* percentages only have reagent contamination DNA to compete with during the PCR step and not any other sample-associated signals. *S. bongori* percentages in the top 20th percentile on average account for 71% of all reads, which would correspond to a sensitivity limit of ~0.2% of reads per CFU (100/(1,100/(71/29)). However, a threshold of 1%, as previously used[9], can be considered a more reliable cut-off for determining whether a signal should be considered biologically relevant. A threshold of 1% would be indicative of multiple replication events (more than 2) and thus metabolic activity or repeated invasion of the tissue by the respective organism. In addition, a 1% threshold for the 16S rRNA data is comparable with the sensitivity of metagenomics as on average 180 *S. bongori* read pairs were detected with metagenomics (Extended Data Fig. 2a). In contrast to 16S analysis, the *S. bongori* spike has no meaningful effect on quantification in metagenomics as microorganisms only represent a very small fraction of the total amount of reads (the vast majority of reads are human). Hence, 6 CFUs are required on average per metagenomics read pair and 6 CFUs would result in a signal of approximately 1% of 16S amplicon reads in cohort 2 using the Qiagen kit.

Nested PCR. We developed a nested PCR assay to sensitively detect the *S. agalactiae sip* gene. In total, 276 placental DNA samples (isolated with the Qiagen kit as described above) were used of which 226 had no (0%) *S. agalactiae* reads detected by 16S rRNA gene sequencing, while *S. agalactiae* reads were detected in 50 samples (range 0.002–63.37% of 16S rRNA reads). The first-round PCR was performed using the DreamTaq PCR Master Mix (2×) (K1071; Thermo Fisher Scientific) and the following primers for the *sip* gene at a final concentration of 0.5 μM: forward 5′-TGAAAATGAATAAAAAGGTACTATTGACAT-3′ and reverse 5′-AAGCTGGCGCAGAAGAATA-3′. Amplification was performed in 50-μl aliquots and using 500 ng of placental DNA per reaction. Genomic *S. agalactiae* DNA (ATCC BAA-611DQ) was used as positive control at 20 or 2 copies per reaction. One reaction was set up with water instead of gDNA as negative control. The PCR amplification profile had an initial step of 95 °C for 3 min followed by 15 cycles of 95 °C (30 s), 48 °C (30 s), and 72 °C (60 s). After completion of cycling, the reactions were incubated for 3 min at 72 °C. The second-round qPCR was performed using the TaqMan Multiplex Master Mix (4461882; Thermo Fisher Scientific) and two TaqMan Assays (Thermo Fisher Scientific): Ba04646276_s1 (Gene Symbol: SIP; Dye Label, Assay Concentration: FAM-MGB, 20×) at a final 1× concentration; RNase P TaqMan assay (ABY dye/QSY probe Thermo Fisher Scientific 4485714) at a final 0.5× concentration, added as a positive control for the

human DNA. In each well, 6 μl of the first-round PCR (or water in the no template control/blank wells) was used as the reaction substrate in a total volume of 15 μl. The PCR amplification profile had an initial step of 95 °C for 20 s followed by 40 cycles of 95 °C (5 s) and 60 °C (20 s).

Statistics. The inter-rater agreement kappa scores[35] and *P* values were computed by DAG_Stat[36]. Comparison of cases and controls was performed using multivariable logistic regression, with conditional logistic regression employed for paired comparisons, using Stata v.15.1 (Statacorp). Other statistical calculations were performed in GraphPad Prism 7 (GraphPad Software). PCAs were performed with the prcomp function from the R package in RStudio (v.0.99.902) with all settings, where applicable, set to 'true'. As the effect size was not known in advance, we performed power calculations with varying prevalence and effect sizes (odds ratio) for 100 case–control pairs (pre-eclampsia and growth restriction) used in the 16S rRNA amplicon sequencing study. These showed that a 5% prevalence in controls and OR = 5 gives 82% power to detect the signal at significance level 0.05. The bioinformatic analysis and the setting of the minimum detection thresholds were performed in a blinded fashion in respect to adverse pregnancy outcome status. All reported *P* values are two-sided except for concordance calculations, as indicated. The experiments were not randomized, and investigators were not blinded to allocation during experiments and outcome assessment unless described otherwise.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The 16S rRNA gene sequencing datasets generated and analysed in this study are publicly available under European Nucleotide Archive (ENA) accession number ERP109246. The metagenomics datasets, which primarily contain human sequences, are available with managed access in the European Genome-phenome Archive (EGA) accession number EGAD00001004198.

20. Pasupathy, D. et al. Study protocol. A prospective cohort study of unselected primiparous women: the pregnancy outcome prediction study. *BMC Pregnancy Childbirth* **8**, 51 (2008).
21. Gardosi, J., Mongelli, M., Wilcox, M. & Chang, A. An adjustable fetal weight standard. *Ultrasound Obstet. Gynecol.* **6**, 168–174 (1995).
22. American College of Obstetricians and Gynecologists & Task Force on Hypertension in Pregnancy. Report of the American College of Obstetricians and Gynecologists' Task Force on Hypertension in Pregnancy. *Obstet. Gynecol.* **122**, 1122–1131 (2013).
23. Lager, S. et al. Detecting eukaryotic microbiota with single-cell sensitivity in human tissue. *Microbiome* **6**, 151 (2018).
24. Walker, A. W. et al. 16S rRNA gene-based profiling of the human infant gut microbiota is strongly influenced by sample processing and PCR primer choice. *Microbiome* **3**, 26 (2015).
25. Wood, D. E. & Salzberg, S. L. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol.* **15**, R46 (2014).
26. Nurk, S. et al. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. *J. Comput. Biol.* **20**, 714–737 (2013).
27. Johnson, M. et al. NCBI BLAST: a better web interface. *Nucleic Acids Res.* **36**, W5–W9 (2008).
28. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
29. Carver, T., Harris, S. R., Berriman, M., Parkhill, J. & McQuillan, J. A. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics* **28**, 464–469 (2012).
30. Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. & Schloss, P. D. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl. Environ. Microbiol.* **79**, 5112–5120 (2013).
31. Eren, A. M. et al. Oligotyping: differentiating between closely related microbial taxa using 16S rRNA gene data. *Methods Ecol. Evol.* **4**, 1111–1119 (2013).
32. Schmieder, R. & Edwards, R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **27**, 863–864 (2011).
33. Quast, C. et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2013).
34. Ludwig, W. et al. ARB: a software environment for sequence data. *Nucleic Acids Res.* **32**, 1363–1371 (2004).
35. Viera, A. J. & Garrett, J. M. Understanding interobserver agreement: the kappa statistic. *Fam. Med.* **37**, 360–363 (2005).
36. Mackinnon, A. A spreadsheet for the calculation of comprehensive statistics for the assessment of diagnostic tests and inter-rater agreement. *Comput. Biol. Med.* **30**, 127–134 (2000).

Acknowledgements The work was supported by the Medical Research Council (UK; MR/K021133/1) and the National Institute for Health Research (NIHR) Cambridge Biomedical Research Centre (Women's Health theme). We thank L. Bibby, S. Ranawaka, K. Holmes, J. Gill, R. Millar and L. Sánchez Busó for technical assistance during the study. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

Author contributions G.C.S.S., D.S.C.-J., J.P. and S.J.P. conceived the experiments. G.C.S.S., D.S.C.-J., J.P., S.J.P. and S.L. designed the experiments. S.L. and M.C.d.G. optimized the experimental approach. S.L. and F.G. performed the experiments. M.C.d.G. analysed all of the sequencing data. U.S. matched cases and controls, performed statistical analyses and provided logistical support for patient and sample metadata. E.C. managed sample collection and processing and the biobank in which all sample were stored. All authors contributed in writing the manuscript and approved the final version.

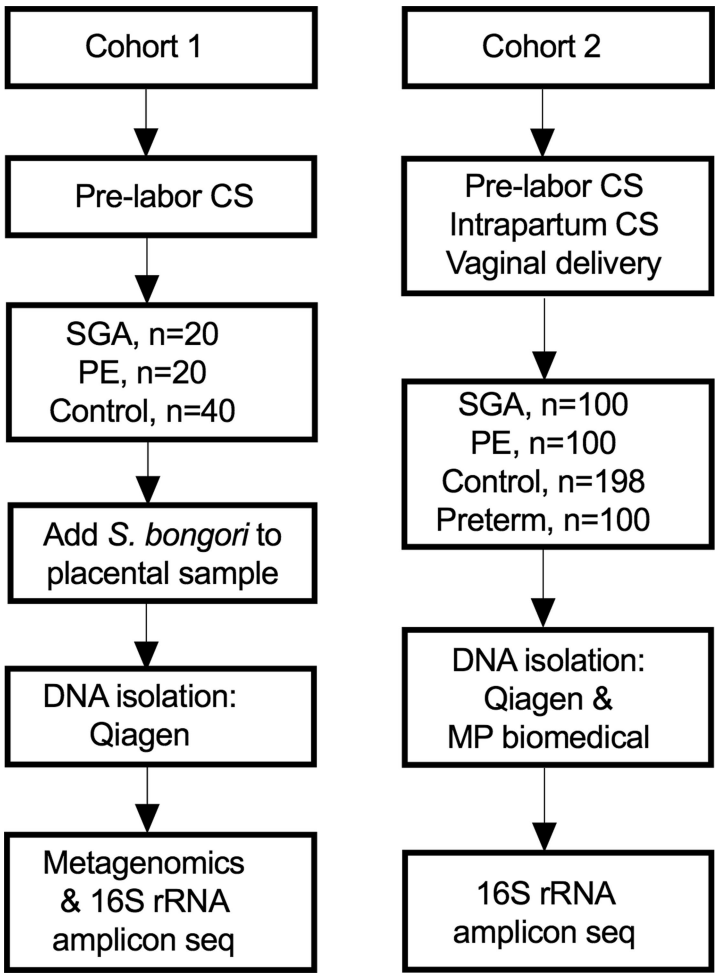
Competing interests J.P. reports grants from Pfizer, personal fees from Next Gen Diagnostics, outside the submitted work; S.J.P. reports personal fees from Specific, personal fees from Next Gen Diagnostics, outside the submitted work; D.S.C.-J. reports grants from GlaxoSmithKline Research and Development, outside the submitted work and non-financial support from Roche Diagnostics, outside the submitted work; G.C.S.S. reports grants and personal fees from GlaxoSmithKline Research and Development, personal fees and non-financial support from Roche Diagnostics, outside the submitted work; D.S.C.-J. and G.C.S.S. report grants from Sera Prognostics, non-financial support from Illumina, outside the submitted work. M.C.d.G., S.L., U.S., F.G. and E.C. have nothing to disclose.

Additional information

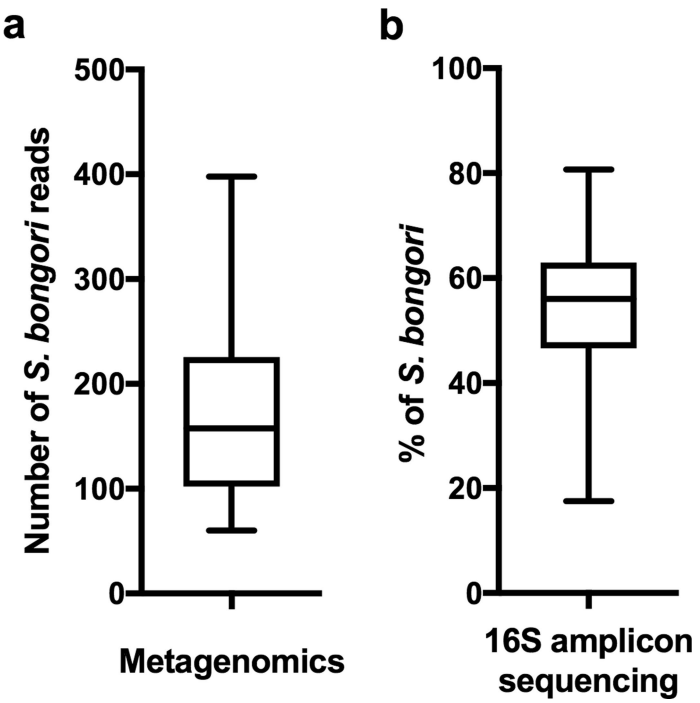
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1451-5.

Correspondence and requests for materials should be addressed to J.P. or G.C.S.S.

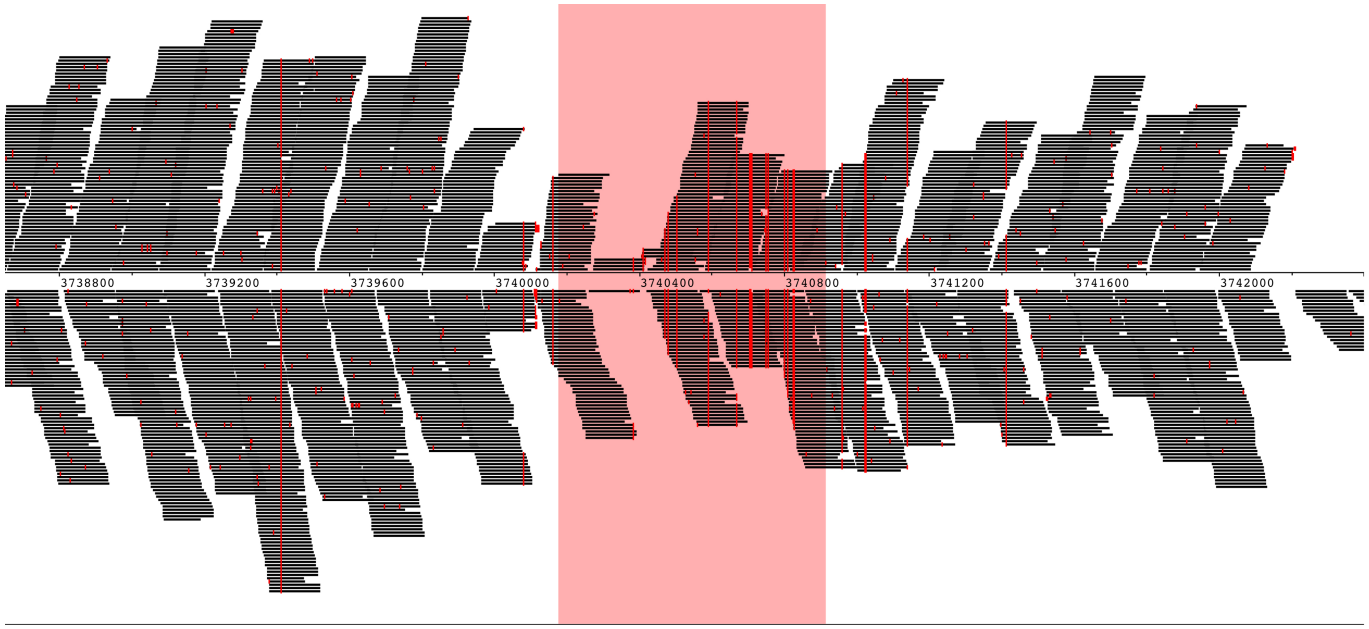
Peer review information *Nature* thanks David N. Fredricks and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



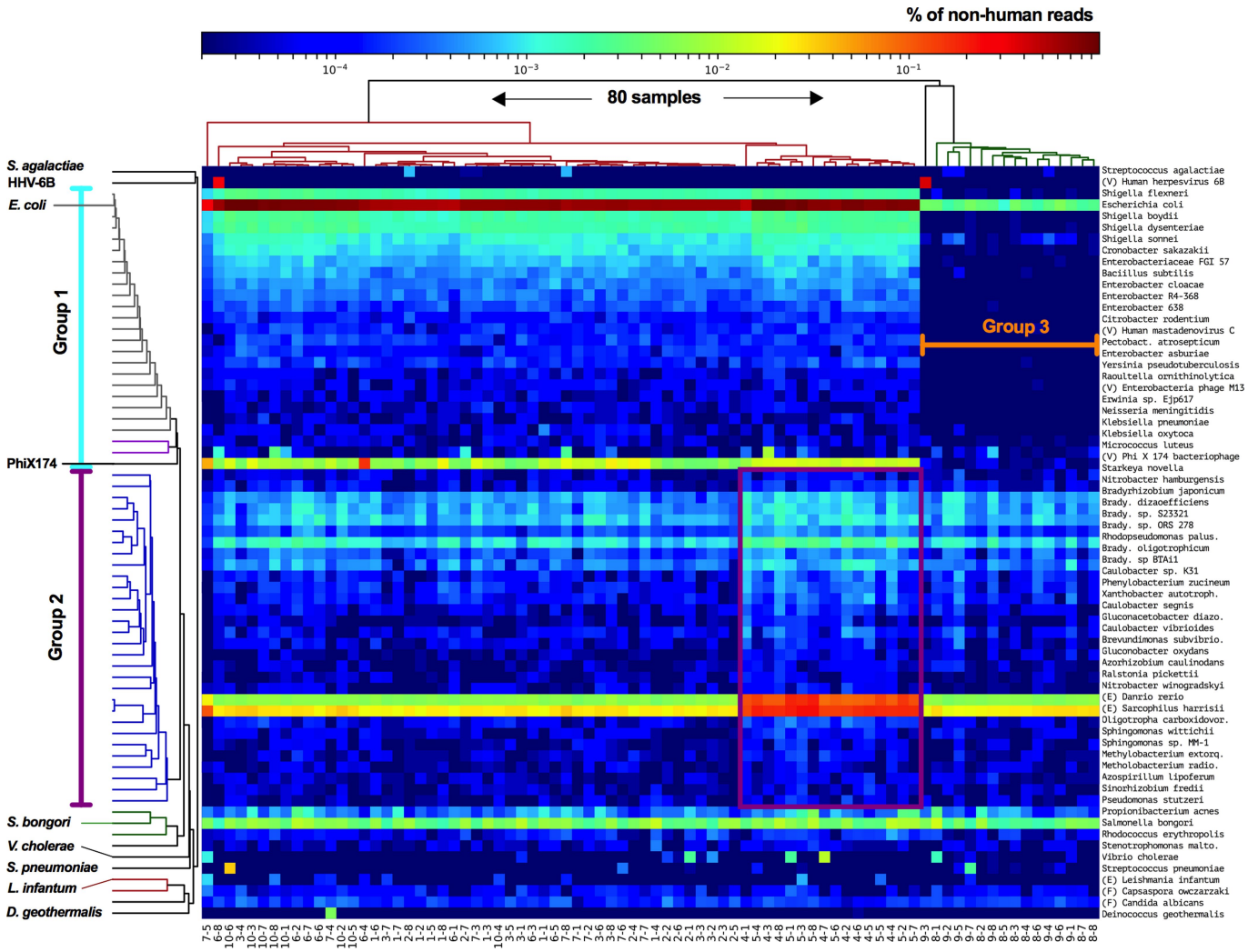
Extended Data Fig. 1 | Two cohorts of placental samples were analysed. Cohort 1 ($n = 80$) contained only samples from pre-labour Caesarean section (CS) deliveries and *S. bongori* was added to the samples before DNA isolation as a positive control. Samples in cohort 1 were analysed by both metagenomics and16S rRNA amplicon sequencing. Cohort 2 ($n = 498$) contained placental samples from Caesarean section and vaginal deliveries. DNA was isolated twice from each placental sample with two different DNA extraction kits. Samples were analysed by 16S rRNA amplicon sequencing. Pre-eclampsia (PE) was defined using The American College of Obstetricians and Gynaecologists (ACOG) 2013 definition. Small for gestational age (SGA) was defined as a birth weight less than the fifth percentile using a customized reference. Preterm denotes birth before 37 weeks gestation.



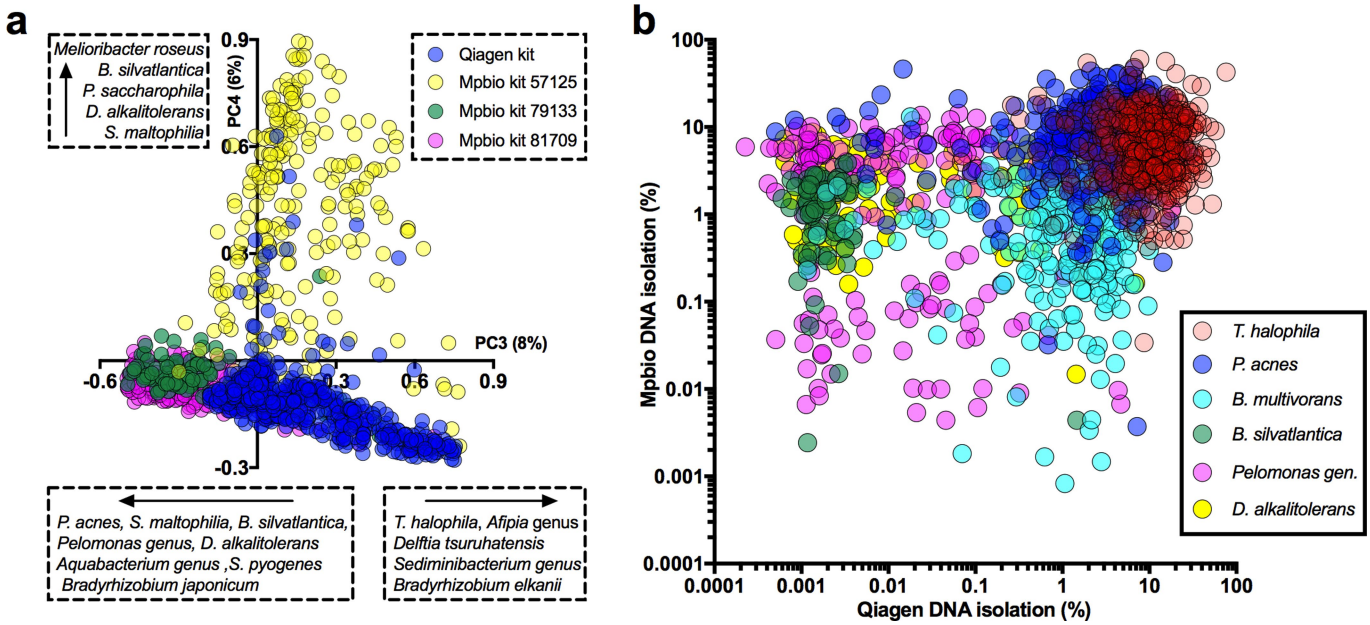
Extended Data Fig. 2 | Positive control experiment comparison between metagenomics and 16S amplicon sequencing. a, **b**, Adding approximately 1,100 CFUs of *S. bongori* to the placental tissue before DNA isolation resulted in an average of 180 reads (s.d. 90 reads) by metagenomic sequencing ($n = 80$) (**a**) or on average of 54% of all 16S rRNA amplicon sequencing reads (approximately 33,000 reads) being identified as *S. bongori* (s.d. 13%; $n = 79$) (**b**). Box represents the interquartile range; whiskers represent the maximum and minimum values; centre lines denote the median.



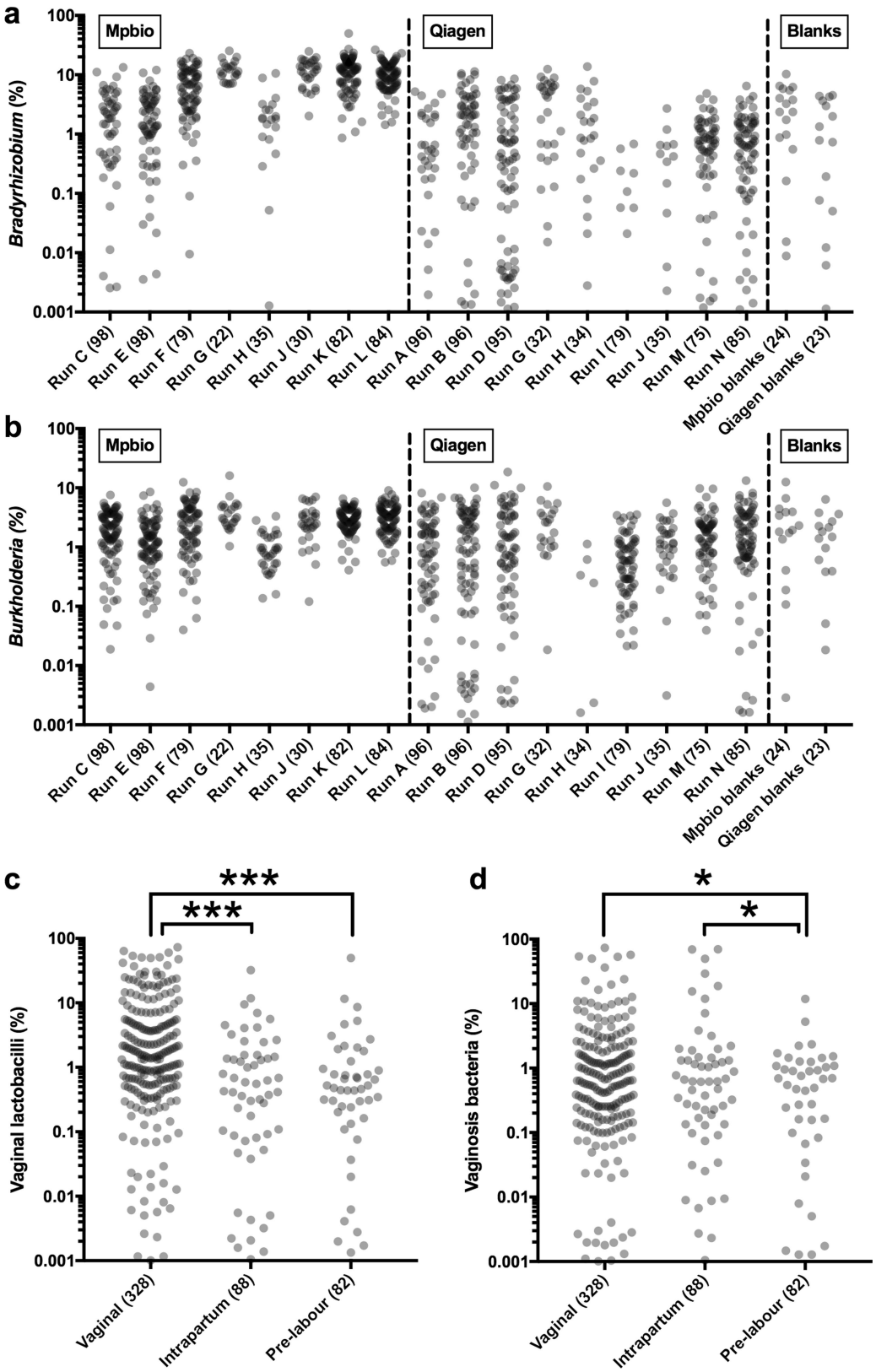
Extended Data Fig. 3 | Strain analysis of *E. coli* reads found by metagenomics. All reads identified in all 80 samples by Kraken[25] as *E. coli* were extracted and mapped together against the closest *E. coli* reference genome (GenBank: CP02409.1). Single nucleotide polymorphisms, shown in red, were consistent for all samples across the genome. Single nucleotide polymorphisms were rare, except in the fimbrial chaperone protein gene (*EcpD*) indicated in light red. Sequence differences that appear as short sporadic red lines represent sequencing errors. Strain variation would have resulted in dashed vertical lines.



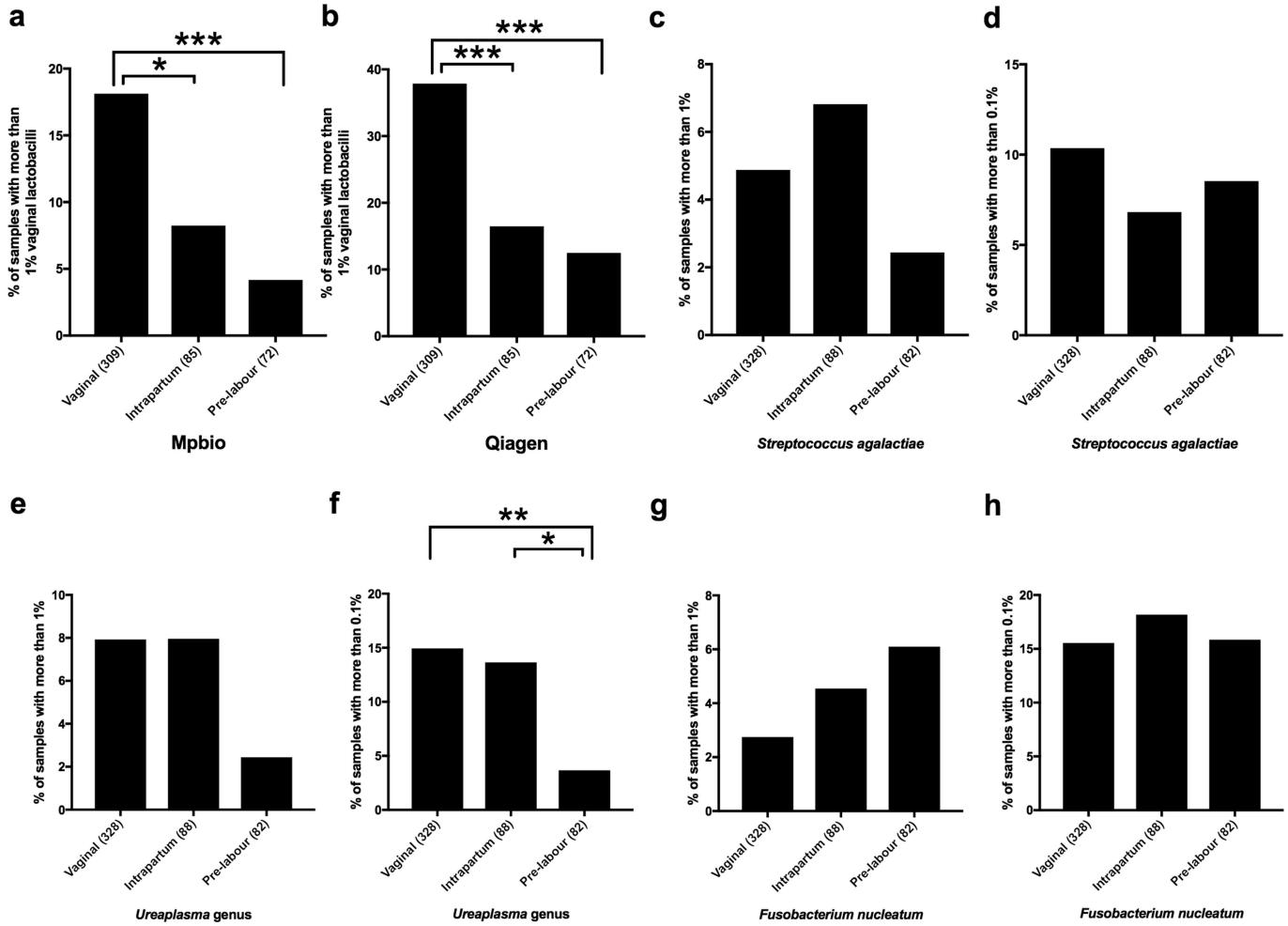
Extended Data Fig. 4 | Detailed heat map metagenomic data. Heat map showing the abundance of all non-human reads as detected by metagenomics. Human reads remaining after filtering (89.8%; s.d. 1.5%) are not shown for scaling purposes. Most taxa (shown on the right) are found in higher abundance within groups 1 and/or 2 (indicated on the left with light blue and purple, respectively). The purple box highlights the samples and species associated with group 2. The lane ID of each sample is represented by the first number (*x* axis). All samples from lanes 4 and 5 form group 2, and all samples from lanes 8 and 9 form group 3 (see Fig. 1a, b).



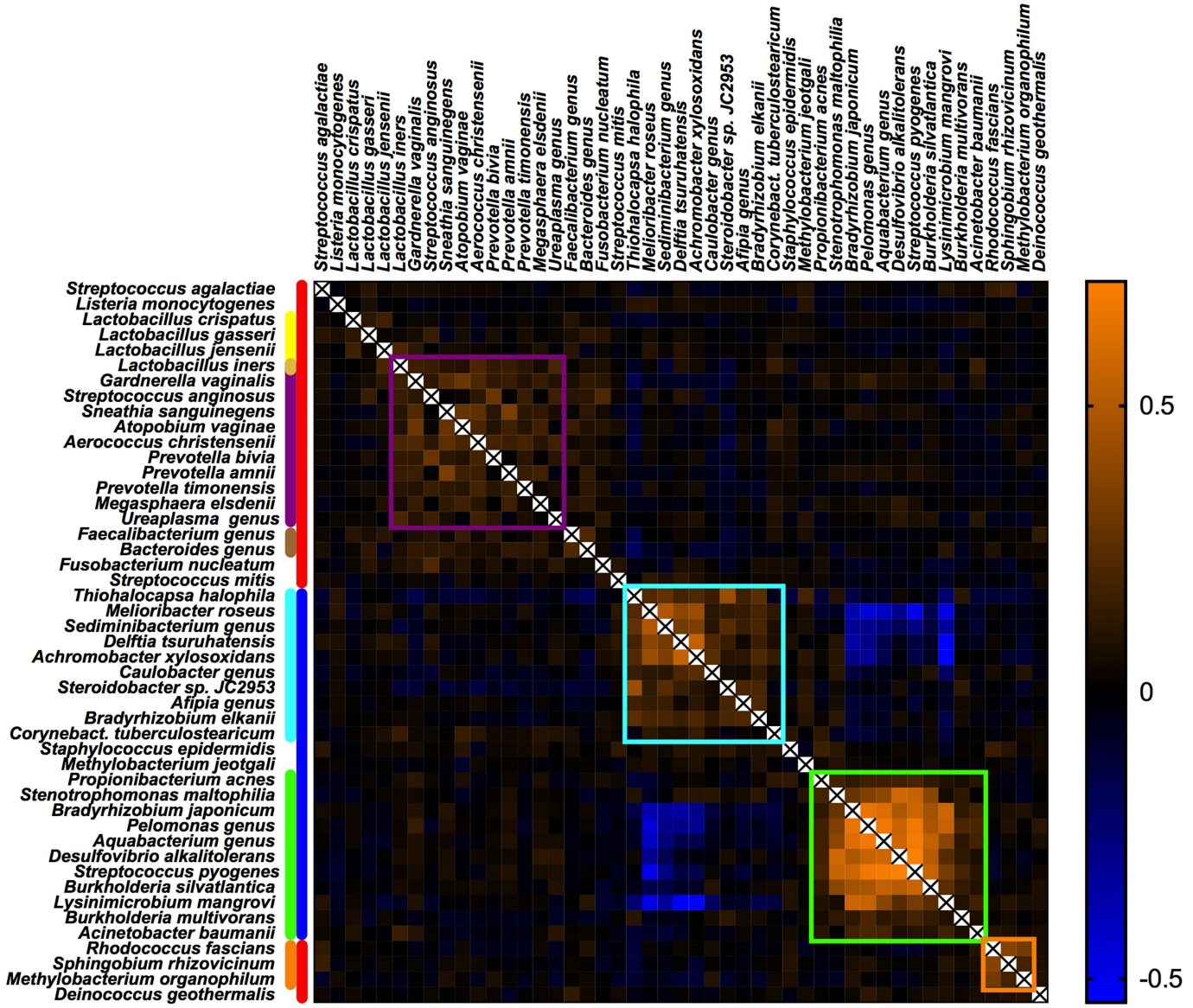
Extended Data Fig. 5 | Species associated with batch effects visualized by PCA also do not show signal reproducibility. a, PCAs of selections of samples from cohort 2 (16S), or of all cohort 2 samples as shown here, allows for the identification of batch effects and allows for the identification of contaminating species associated with the use of specific DNA isolation methods, kits and/or other reagents. An analysis of all samples shows that principal components 3 (*x* axis) and 4 (*y* axis) are strongly correlated with the use of Qiagen or specific Mpbio DNA isolation kits. **b**, Examples of bacteria detected in high abundance and frequency when processed with the Qiagen (*x* axis) and/or Mpbio (*y* axis) DNA isolation kits. Patterns that lack positive correlation (compare with Fig. 2a) demonstrate that signals are not sample- but batch-associated.



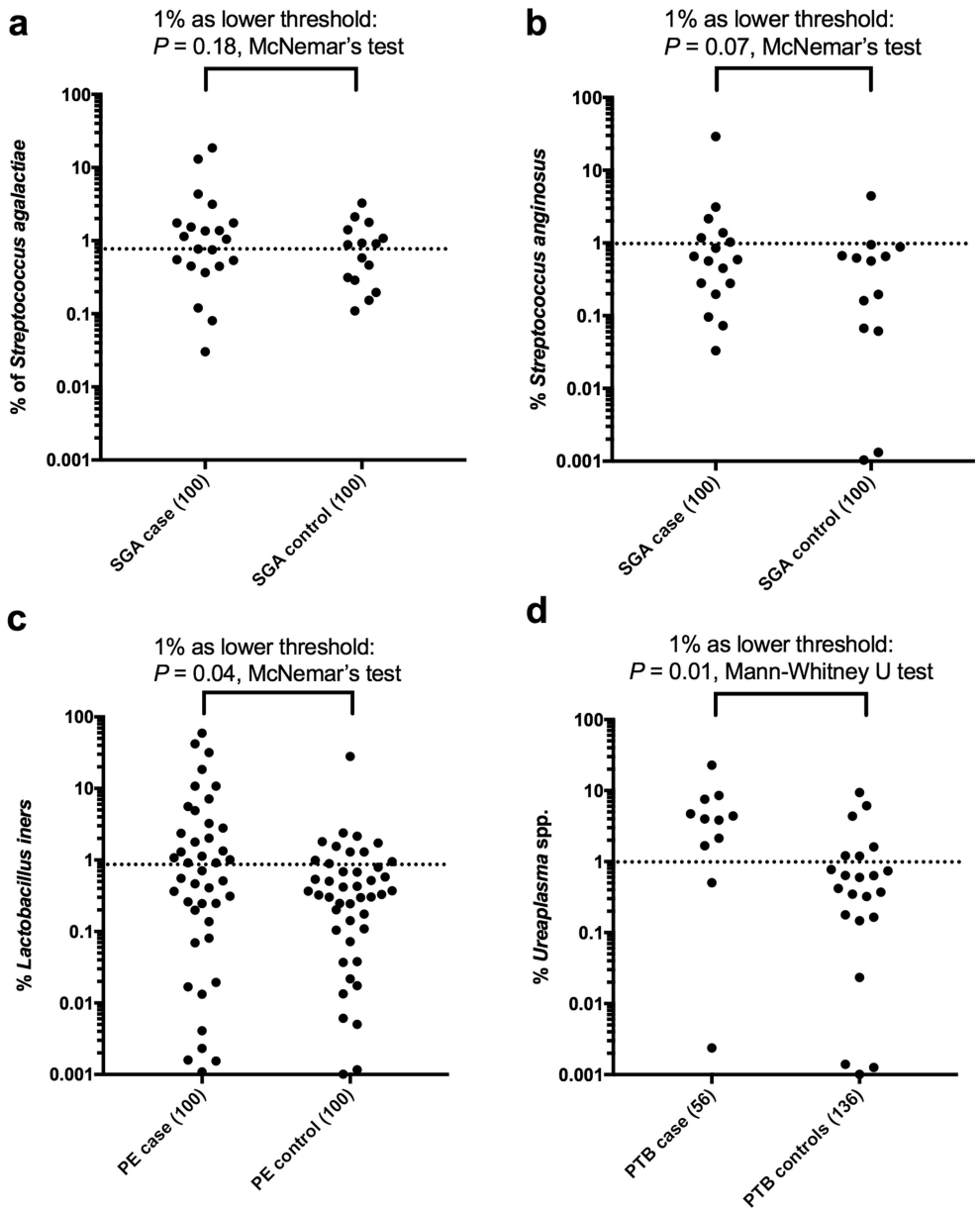
Extended Data Fig. 6 | Scatterplot representations of the abundance of *Bradyrhizobium*, *Burkholderia*, vaginal lactobacilli and vaginosis bacteria during 16S amplicon sequencing. a, **b**, The abundance of *Bradyrhizobium* (**a**) or *Burkholderia* (**b**) with respect to sequencing run batch effects during 16S amplicon sequencing. Numbers in parentheses indicate the number of samples sequenced in a given run. Values of zero are not shown on the logarithmic axis. **c**, **d**, The abundance of vaginal lactobacilli (**c**) and vaginosis bacteria (**d**) with respect to the mode of delivery during 16S amplicon sequencing. $^{*}P < 0.05$, $^{***}P < 0.001$, Mann–Whitney *U*-tests, where values below 1% are regarded as 0% (not biologically relevant).



Extended Data Fig. 7 | Mode of delivery and the detection of bacterial signals. a, **b**, The association of vaginal lactobacilli with the mode of delivery, as determined by the analysis of 466 samples by 16S amplicon sequencing that were successfully sequenced twice using the Mpbio (**a**) and Qiagen (**b**) DNA isolation methods. Comparisons of the Mpbio and Qiagen DNA isolation techniques highlight that the same patterns are observed in the associations with mode of delivery. Comparisons also show that the Qiagen DNA isolation was more sensitive, resulting in twice as many signals above the 1% threshold. **c**–**h**, The association of bacterial groups with mode of delivery. Analyses were performed using all 498 placental samples with the highest value of either DNA isolation method for each bacterial group per sample. **c**, **d**, *S. agalactiae* was not associated with the mode of delivery irrespective of whether a 1% threshold was used (the minimum percentage considered to be potentially ecologically relevant) (**c**) or a 0.1% threshold was used (the 16S detection limit, relevant for detecting traces of contamination during delivery) (**d**). **e**, **f**, The *Ureaplasma* genus was significantly associated with the mode of delivery using the 0.1% threshold, similar to Fig. 2c, which describes the combination of all vaginosis-associated bacteria. **g**, **h**, *F. nucleatum* was not associated with the mode of delivery, irrespective of whether a 1% (**g**) or 0.1% (**h**) threshold was used. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, Mann–Whitney *U*-tests.



Extended Data Fig. 8 | Heat map of Spearman's rho correlation coefficients of bacterial signals as found by 16S rRNA amplicon sequencing. Sample-associated signals (red bar), are typically identified by increased kappa scores, as shown in Supplementary Table 4. Reagent contaminants are indicated by a blue bar. Vaginosis-associated bacteria (purple bar) show positive correlations (purple square) with each other, *Lactobacillus iners* and faecal bacteria (brown bar). Lactobacilli (yellow bar) show limited positive correlation with faecal bacteria. Reagent contaminants mainly associated with the Qiagen (light blue) or the Mpbio (green) kit form distinct clusters. Species that are strongly associated with sample collection contamination in 2012–2013 are indicated in orange. For each species the highest value (percentage) found using either the Qiagen or the Mpbio DNA isolation kit, was used as input (using all 498 samples).



Extended Data Fig. 9 | Bacterial signals and adverse pregnancy outcome. a–d, Scatterplot representations of the non-significant associations of *S. agalactiae* with SGA (**a**), *S. anginosus* with SGA (**b**), and of the significant associations of *L. iners* with pre-eclampsia (**c**), and *Ureaplasma* with PTB (**d**). Samples with 0% signal are not shown on the logarithmic scale. Signals above 1% (dotted line) are regarded as positive for use in McNemar's test (**a–c**), and signals below 1% are considered as negative. The Mann–Whitney *U*-test was used for unpaired samples in **d**.

nature research

Corresponding author(s): Gordon C. S. Smith

Last updated by author(s): May 30, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	The only software used to collect data was the standard MiSeq and HiSeq (Illumina) sequencing machine software and the quantitative PCR machine software (QuantStudio 6 Flex system, ThermoFisher Scientific).
Data analysis	KneadData (v0.6.1), Kraken (v0.10.6), Mothur (v1.40.5), PRINSEQ-lite (v0.20.3), oligotyping (v2.1) , ARB (v5.5-org-9167), DAG_Stat, Stata (v15.1), R package RStudio (v0.99.902), Past3 (v3.14), Prism 7 (v7.0c), Spades (v3.11.0), BWA (v0.7.17-r1188), Artemis (v.16.0.0), BLASTN (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and custom script was used to extract reads identified of a particular group of interest identified by Kraken (Supplemental Information).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The 16S rRNA gene sequencing datasets utilized in this study are publicly available under European Nucleotide Archive (ENA) accession no. ERP109246. The metagenomics data sets, which primarily contain human sequences, are available in the European Genome-phenome Archive (EGA) with managed access (EGAD00001004197).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	A power calculation was performed during the planning phase of the Pregnancy Outcome Prediction (POP) study and it is described in Pasupathy et al (BMC Pregnancy and Childbirth 2008 PMID 19019223). In brief, the sensitivity of different models for a given screen positive rate was quantified by 95% confidence intervals. The calculations indicated that the study was likely to provide reasonably precise estimates of sensitivity for conditions with a 3% incidence, such as severe SGA. The use of a nested case-control design with a 1:1 matching of cases and controls on key maternal characteristics was also planned in advance in the context of very expensive or labor intensive methodologies (Pasupathy et al). For the 16S rRNA amplicon sequencing study we used 100 matched cases and controls for both pre-eclampsia and growth restriction (ie 200 samples in total). As the effect size was not known in advance we performed power calculations with varying prevalence and effect sizes (OR) for 100 case-control pairs. These showed that a 5% prevalence in controls and OR=5 gives 82% power to detect the signal at significance level 0.05.
Data exclusions	A total of 4512 women with a viable singleton pregnancy were recruited to the POP study. The only clinical exclusion criterion was multiple pregnancy.
Replication	Reproducibility of signals was confirmed by analyzing samples both by metagenomic and 16S rRNA amplicon analysis (cohort 1) and by analysing each sample from cohort 2 twice by 16S rRNA amplicon sequencing using 2 different DNA isolation methods. A large part of the manuscript is about proving the reproducibility of signals in order to show which signals are real and which ones are spurious
Randomization	The POP study is a prospective cohort study of nulliparous women attending the Rosie Hospital (Cambridge, UK) for their dating ultrasound scan. All eligible participants were included. For the purpose of the experimental projects described in this manuscript, participants were allocated into groups based on pregnancy outcome (details in Methods and Supplementary information). Outcome data were ascertained by review of each woman's paper case record by research midwives and by record linkage to clinical electronic databases. Paired cases and controls were always processed together and sequenced in the same run.
Blinding	All the aspects of the POP study were conducted blind: the results of the research ultrasound scans and the biochemical marker data were not revealed to the clinicians, patients and researchers performing the downstream experiments. Data were unblinded only at the statistical analysis stage. Specifically, all of the bioinformatic analysis of 16S rRNA amplicon data and the metagenomic data was performed in a blinded fashion. Reagent contamination recognition was also performed prior to unblinding. Finally, a statistical analysis plan was written prior to unblinding for the analysis of Streptococcus agalactiae, the only bacterial signal that passed all quality checks for being a genuine and possibly important. All other bacterial analyses (done for all the other bacteria) should be considered exploratory.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	Samples were from the Pregnancy Outcome Prediction (POP) study. In the whole POP study population (n=4212), the median age, height and BMI (IQR) were 30.3 (26.8 to 33.4) years, 165 (161 to 169) cm, 24.1 (21.8 to 27.3) kg/m2, respectively, and 13% of the women were smokers at recruitment. Detailed characteristics of women whose samples were selected for sequencing in this study are given in Extended Data Tables 1 and 2. In brief, the median maternal age varied between 29.7 and 30.9 years between the groups of 100 cases or controls (Extended Data Table 2). The median height was similar (164-165 cm) between the groups. The median BMI was highest in the PE cases (25.7 kg/m2) and otherwise varied between 24.1 and 25.0 kg/m2 between the groups. The prevalence of smoking at booking varied the most; it was 28% in the SGA group and 7% among the controls of PE cases.
Recruitment	Samples were from the Pregnancy Outcome Prediction (POP) study. Nulliparous women with a viable singleton pregnancy who attended their dating ultrasound scan at the Rosie Hospital (Cambridge, UK) between 14 January 2008 and 31 July 2012 were eligible (n=8028), and 4512 (56%) of them provided an informed consent and were recruited. The recruited and non-recruited women were broadly comparable, although according to the hospital record data the women who were recruited were slightly older, more often of white ethnic origin and less likely to smoke. In addition, women were excluded because they delivered elsewhere (n=233) or withdrew their consent (n=67). The cohort of 4212 women used for the sample selection in the present study can be regarded as fairly well representative of the eligible population. See Sovio et al Lancet 2015 PMID 26360240 and Gaccioli et al Placenta 2017 PMCID PMC5701771 for a complete description.
Ethics oversight	The Pregnancy Outcome Prediction study was approved by the Cambridgeshire 2 Research Ethics Committee (reference number 07/H0308/163).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

ARTICLE

https://doi.org/10.1038/s41586-019-1406-x

Activation of PDGF pathway links *LMNA* mutation to dilated cardiomyopathy

Jaecheol Lee[1,2,3,4,12]*, Vittavat Termglinchan[1,2,3,12], Sebastian Diecke[5,6,7,12], Ilanit Itzhaki[1,2,3], Chi Keung Lam[1,2,3], Priyanka Garg[1,2,3], Edward Lau[1,2,3], Matthew Greenhaw[8], Timon Seeger[1,2,3], Haodi Wu[1,2,3], Joe Z. Zhang[1,2,3], Xingqi Chen[9], Isaac Perea Gil[1,8], Mohamed Ameen[1,2,3], Karim Sallam[1,2,3], June-Wha Rhee[1,2,3], Jared M. Churko[1,2,3], Rinkal Chaudhary[1,2,3], Tony Chour[1,2,3], Paul J. Wang[2], Michael P. Snyder[1,10], Howard Y. Chang[9,11], Ioannis Karakikes[1,8]* & Joseph C. Wu[1,2,3]*

Lamin A/C (*LMNA*) is one of the most frequently mutated genes associated with dilated cardiomyopathy (DCM). DCM related to mutations in *LMNA* is a common inherited cardiomyopathy that is associated with systolic dysfunction and cardiac arrhythmias. Here we modelled the LMNA-related DCM in vitro using patient-specific induced pluripotent stem cell-derived cardiomyocytes (iPSC-CMs). Electrophysiological studies showed that the mutant iPSC-CMs displayed aberrant calcium homeostasis that led to arrhythmias at the single-cell level. Mechanistically, we show that the platelet-derived growth factor (PDGF) signalling pathway is activated in mutant iPSC-CMs compared to isogenic control iPSC-CMs. Conversely, pharmacological and molecular inhibition of the PDGF signalling pathway ameliorated the arrhythmic phenotypes of mutant iPSC-CMs in vitro. Taken together, our findings suggest that the activation of the PDGF pathway contributes to the pathogenesis of LMNA-related DCM and point to PDGF receptor-β (PDGFRB) as a potential therapeutic target.

DCM associated with mutations in *LMNA* (LMNA-related DCM) is an autosomal dominant disorder caused by mutations in the gene that encodes the lamin A/C proteins that constitute the major component of the nuclear envelope[1–3]. LMNA-related DCM accounts for 5–10% of cases of DCM and has an age-related penetrance with a typical onset[4,5] between the ages of 30 and 40. In contrast to most other forms of familial DCM, sudden cardiac death may be the first manifestation of LMNA-related DCM even in the absence of systolic dysfunction, owing to malignant arrhythmias such as ventricular tachycardia and fibrillation[4–6]. However, the precise mechanisms that link the mutations in *LMNA* to increased arrhythmogenicity are unknown.

Modelling LMNA-related DCM with iPSC-CMs in vitro

We recruited a large family cohort, members of which carry a frameshift mutation in *LMNA* that leads to the early termination of translation (348–349insG; K117fs) (Extended Data Fig. 1a–c). Three of the carriers (III-1, III-3 and III-9) presented with atrial fibrillation that progressed to atrioventricular block, ventricular tachycardia (Extended Data Fig. 1d, e) and DCM.

We generated multiple patient-specific iPSC lines using non-integrating reprogramming methods[7,8] and derived iPSC-CMs using a chemically defined protocol[8–10] to examine the electrophysiological properties at the single-cell level. We found that the *LMNA*-mutant iPSC-CMs (III-3, III-9, III-15 and III-17) exhibited proarrhythmic activity in both atrial- and ventricular-like iPSC-CMs compared to healthy controls (IV-1 and IV-2) (Fig. 1a and Extended Data Fig. 1f, g). Taken together, these data demonstrate that patient-specific iPSC-CMs recapitulate the disease phenotype associated with LMNA-related DCM in vitro.

Next, we generated a panel of isogenic lines that differed only in this mutation using the iPSC line derived from patient III-3 (who carried one wild-type and one mutant allele (WT/MUT)) through TALEN-mediated genome editing[11,12]. Specifically, we corrected the *LMNA* mutation to the wild-type allele in the iPSCs (WT/cor-WT), inserted the K117fs mutation in the wild-type allele (ins-MUT/MUT) and generated a knockout iPSC line by targeting the start codon[11] (ATG site) of the wild-type allele (del-KO/MUT) (Fig. 1b and Extended Data Fig. 2a–c). We also introduced the K117fs mutation in the healthy control iPSC line (patient IV-1, who carried two wild-type alleles (WT/WT)) to generate a heterozygous mutant iPSC line (WT/ins-MUT). We generated iPSC-CMs from the isogenic lines and observed that the targeted gene correction rescued the electrophysiological abnormalities in WT/cor-WT-derived iPSC-CMs compared to parental WT/MUT, genome-edited ins-MUT/MUT and del-KO/MUT iPSC-CMs (Fig. 1c–g). As expected, the insertion of the K117fs mutation in the line derived from the healthy control individual (WT/ins-MUT) induced arrhythmias (Extended Data Fig. 2d–g). Together, these data suggest that LMNA K117fs is a pathogenic mutation that causes LMNA-related DCM.

As homeostasis of calcium ions (Ca^{2+}) is critical for excitation–contraction coupling in the heart[13,14], we analysed the intracellular Ca^{2+}-handling properties of the isogenic iPSC-CMs. Abnormal Ca^{2+} transients were observed in K117fs iPSC-CMs, whereas the control iPSC-CMs exhibited uniform Ca^{2+} transients (Fig. 2a). Furthermore, WT/ins-MUT iPSC-CMs displayed abnormal Ca^{2+} transients when compared to the isogenic WT/WT iPSC-CMs (Fig. 2b). Next, we recorded the calcium transient in the presence of tetrodotoxin, a sodium channel blocker, to inhibit any beating initiated at the plasma membrane[15,16]. We observed spontaneous Ca^{2+} cycling at very low extracellular Ca^{2+} levels in WT/MUT iPSC-CMs in contrast to the

[1]Stanford Cardiovascular Institute, Stanford University, Stanford, CA, USA. [2]Department of Medicine, Division of Cardiovascular Medicine, Stanford University, Stanford, CA, USA. [3]Institute for Stem Cell Biology and Regenerative Medicine, Stanford University, Stanford, CA, USA. [4]School of Pharmacy, Sungkyunkwan University, Suwon, South Korea. [5]Berlin Institute of Health, Berlin, Germany. [6]Max Delbrueck Center, Berlin, Germany. [7]DZHK (German Centre for Cardiovascular Research), partner site Berlin, Berlin, Germany. [8]Department of Cardiothoracic Surgery, Stanford University School of Medicine, Stanford, CA, USA. [9]Center for Personal Dynamic Regulomes, Stanford University, Stanford, CA, USA. [10]Department of Genetics, Stanford University School of Medicine, Stanford, CA, USA. [11]Howard Hughes Medical Institute, Stanford University, Stanford, CA, USA. [12]These authors contributed equally: Jaecheol Lee, Vittavat Termglinchan, Sebastian Diecke. *e-mail: jaecheol@skku.edu; ioannis1@stanford.edu; joewu@stanford.edu

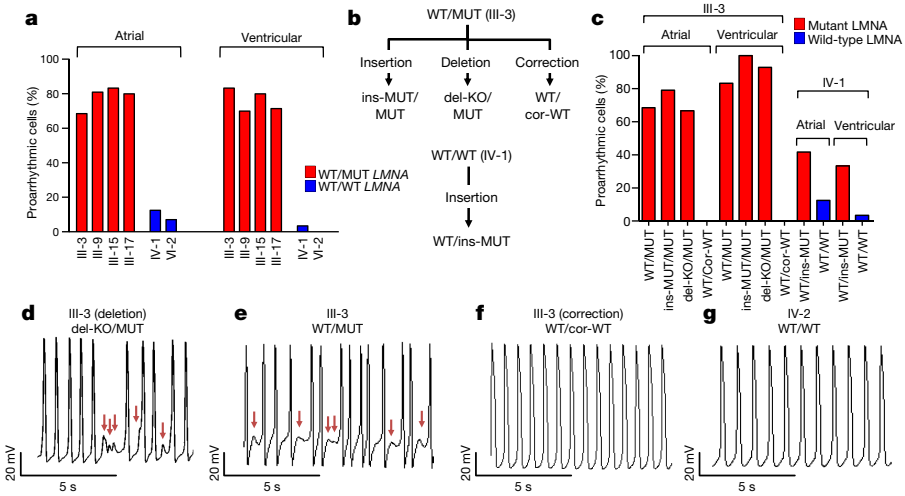


Fig. 1 | The mutation in *LMNA* causes an arrhythmic phenotype in patient-specific iPSC-CMs. a, Quantification of the occurrence of arrhythmias in control and mutant iPSC-CMs. **b**, Schematic view of genome-editing strategy. **c**, Quantification of the occurrence of arrhythmias in isogenic iPSC-CMs. **d**–**g**, Electrophysiological measurements of spontaneous action potentials in parental mutant iPSC-CMs (III-3 WT/MUT), isogenic mutant iPSC-CMs (III-3 del-KO/MUT), isogenic control iPSC-CMs (III-3 WT/cor-WT) and control iPSC-CMs (IV-2 WT/WT). Red arrows denote delayed afterdepolarizations. The experiments were independently repeated three times with similar results.

minimal occurrence of Ca^{2+} activity found in the isogeneic control line (WT/cor-WT), suggesting that abnormal calcium release from the sarcoplasmic reticulum occurred in the WT/MUT iPSC-CMs (Extended Data Fig. 3a–c). Taken together, these findings demonstrated that the dysregulation of Ca^{2+} in the sarcoplasmic reticulum is associated with the electrical abnormalities observed in K117fs iPSC-CMs.

Given that hyperphosphorylation of ryanodine receptor 2 (RYR2) by Ca^{2+}/calmodulin-dependent kinase II (CAMK2) leads to arrhythmias related to delayed afterdepolarizations[17], as documented in the LMNA-mutant iPSC-CMs (Fig. 1d, e), we tested whether the activation of this pathway induces arrhythmias in K117fs iPSC-CMs. Notably, phosphorylated RYR2 (pRYR2) and phosphorylated CAMK2D (pCAMK2D) levels were significantly increased in K117fs iPSC-CMs (WT/MUT, ins-MUT/MUT and del-KO/MUT) compared to the levels found in the isogenic control iPSC-CMs (WT/cor-WT) (Fig. 2c, d). By contrast, expression levels of both *CAMK2D* and *RYR2* mRNA were similar between isogenic control and K117fs iPSC-CMs (Extended Data Fig. 3d–g). When the activation of CAMK2D was inhibited in K117fs iPSC-CMs using KN93, a specific CAMK2D inhibitor, we observed a significant decrease in the levels of pRYR2 and pCAMK2D as well as a significant decrease in abnormal Ca^{2+} transients (22.22%, $n = 81$) compared to K117fs iPSC-CMs treated with vehicle (65.38%, $n = 52$) or the inactive analogue KN92 (65.30%, $n = 49$) (Fig. 2e and Extended Data Fig. 3h–j). Taken together, these data suggest that CAMK2-mediated RYR2 activation causes abnormal Ca^{2+} handling and arrhythmias in K117fs iPSC-CMs.

Lamin A/C haploinsufficiency in mutant iPSC-CMs

Given that abnormalities in nuclear structures are associated with laminopathies[18], we examined the integrity of the nuclear envelope in K117fs iPSC-CMs. Through immunostaining analyses, we demonstrated that K117fs iPSC-CMs display abnormal nuclear structures compared to isogenic controls (Fig. 3a and Extended Data Fig. 4a–c). Notably, the expression of lamin A/C proteins was significantly reduced in K117fs compared to isogenic control iPSC-CMs. Furthermore, the full-length or truncated lamin A/C were not detected in ins-MUT/MUT and del-KO/MUT iPSC-CMs (Fig. 3b, c and Extended Data Fig. 4d–f). These data suggest that the K117fs mutation leads to lamin A/C haploinsufficiency. Furthermore, the total level of *LMNA* mRNA expression was significantly reduced in K117fs compared to isogenic control iPSC-CMs (Fig. 3d and Extended Data Fig. 4g).

Nonsense-mediated mRNA decay (NMD) is a mechanism coupled to translation that selectively degrades mRNAs that contain premature translation-termination codons[19,20]. To investigate whether NMD influences the expression levels of *LMNA* mRNA in K117fs iPSC-CMs, we assessed allele-specific expression of *LMNA* mRNA. We found that 97% and 3% of the total *LMNA* mRNA was expressed by the wild-type and the K117fs allele, respectively, in the K117fs iPSC-CMs (WT/MUT; III-3) (Fig. 3e and Extended Data Fig. 4h). We observed a significant increase in the expression levels of the K117fs allele (18–37%) and the appearance of a 14-kDa band upon inhibition of the NMD pathway in K117fs iPSC-CMs (Fig. 3f and Extended Data Fig. 4i, j). In addition, the 14-kDa band was not detected in the isogenic control line (WT/cor-WT) after NMD inhibition (Extended Data Fig. 4k), which suggests that the truncated lamin A/C is translated from the mutant *LMNA* mRNA. Collectively, these findings indicate that NMD-mediated degradation of mutant *LMNA* mRNA induces lamin A/C haploinsufficiency in K117fs iPSC-CMs.

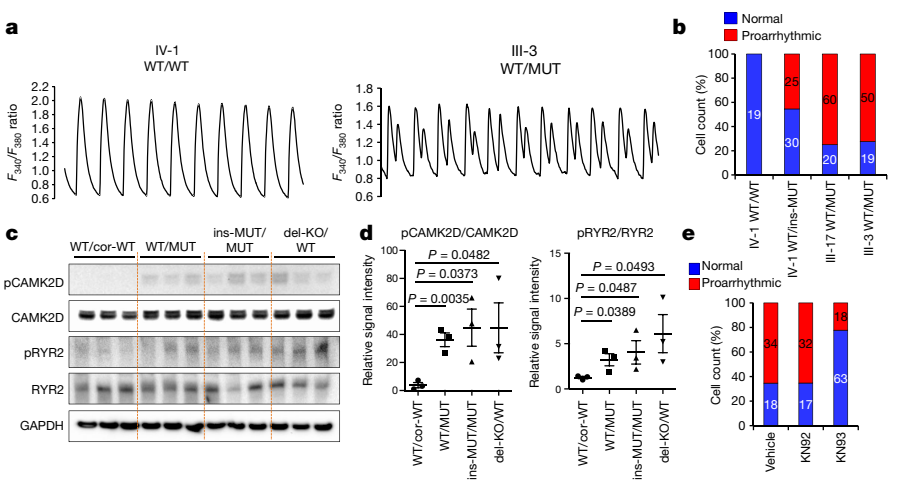


Fig. 2 | Abnormal calcium handling as a cause of the arrhythmic phenotype of *LMNA*-mutant iPSC-CMs. a, Representative Ca^{2+} transients of control and mutant iPSC-CMs. The ratio of the fura-2 AM signal excited at 340 nm and 380 nm is shown (F_{340}/F_{380}). **b**, The percentage of cell count that exhibit arrhythmic waveforms in control and mutant iPSC-CMs. **c**, **d**, Immunoblot analysis of the levels of pRYR2, RYR2, pCAMK2D and CAMK2D in control and mutant iPSC-CMs. Data are mean ± s.e.m.; a two-tailed Student's *t*-test was used to calculate *P* values; $n = 3$. **e**, The percentage of cell count that exhibit arrhythmic waveforms in mutant iPSC-CMs (III-3 WT/MUT) treated with 1 μM of KN92 or KN93 for 24 h. All traces were recorded for 20 s. The Ca^{2+} transients shown in **a** were independently repeated as described in **b** with similar results. The immunoblot data in **c** were independently repeated twice with similar results.

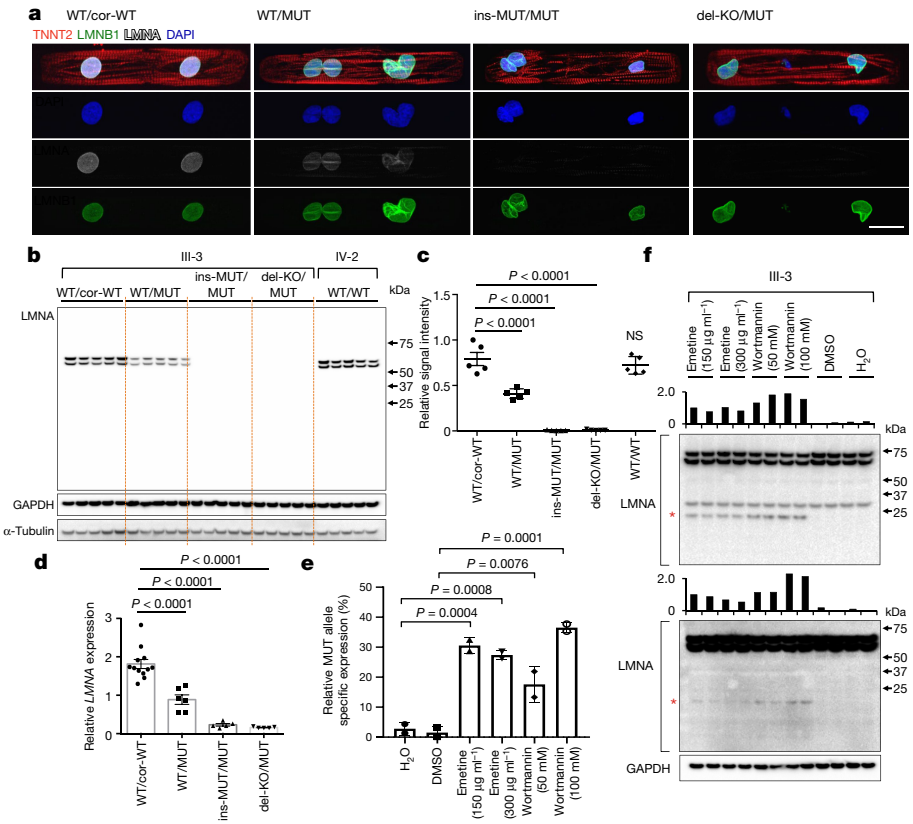


Fig. 3 | NMD pathway-mediated suppression of *LMNA*-mutant mRNA leads to lamin A/C haploinsufficiency in mutant iPSC-CMs. **a**, Representative confocal images of control and mutant lines. Micro-patterned iPSC-CMs were stained with specific antibodies against TNNT2 (red), LMNA (white) and LMNB1 (green). Blue, 4′,6-diamidino-2-phenylindole (DAPI). Scale bar, 20 μm. The immunofluorescence data were independently repeated three times with similar results. **b**, Immunoblot analysis of the levels of lamin A/C in control and mutant iPSC-CMs. The anti-LMNA E-1 antibody was used (lot C1413). GAPDH and α-tubulin are used as loading controls. **c**, Quantification of signal intensity of the LMNA bands shown in **b**. $n = 4$. **d**, Relative mRNA expression of total *LMNA* in control and mutant iPSC-CMs. $n = 12$ (WT/cor-WT), $n = 6$ (WT/MUT), $n = 5$ (ins-MUT/MUT and del-KO/MUT). **e**, Digital droplet PCR analysis of allele-specific expression of *LMNA* in mutant iPSC-CMs treated with emetine (150 or 300 μg ml^{-1} for 6 h) and wortmannin (50 or 100 mM for 6 h). Data are mean ± s.d.; $n = 2$; statistical significance was calculated using a one-way analysis of variance (ANOVA). **f**, Immunoblot analysis of cell lysates from mutant iPSC-CMs treated with emetine and wortmannin. Two different batches of antibodies (E-1, lots A3118 (top) and G1413 (bottom)) were used. Red asterisks indicate the truncated lamin A/C (about 14 kDa in size). The signal intensity of the truncated lamin A/C is shown. The immunoblot data in **b** and **f** were independently repeated twice with similar results. **c**–**e**, Data are mean ± s.e.m.; statistical significance was calculated using one-way ANOVA. NS, not significant.

Increased open chromatin in mutant iPSC-CMs

Lamin A/C interacts with heterochromatin-rich genomic regions at the nuclear envelope called lamin-associated domains (LADs), which have an essential role in the organization of chromatin[21–24]. We therefore postulate that lamin A/C haploinsufficiency could disturb chromatin distribution, leading to aberrant gene expression in K117fs iPSC-CMs. Using an assay for transposase-accessible chromatin with visualization (ATAC-see)[25], we observed that the distribution of open chromatin was biased towards the nuclear periphery in K117fs iPSC-CMs, whereas isogenic control iPSC-CMs showed a uniform distribution throughout the nucleus (Extended Data Fig. 5a–g).

To study whether lamin A/C haploinsufficiency results in an abnormal conformation of open chromatin, we investigated the relationship between LADs and gene activation[26,27] (Extended Data Fig. 6a). We compared the LADs in isogenic iPSC-CMs and grouped these LADs into three categories: loss, overlapping and gain (Fig. 4a and Extended Data Fig. 6b). The genomic coverage, mean LAD length and numbers of LADs were similar in K117fs and isogenic control iPSC-CMs (Fig. 4b–d and Extended Data Fig. 6c–e). Notably, the LADs that showed loss or gain were located in nearby overlapping regions in K117fs iPSC-CMs, which suggests that lamin A/C haploinsufficiency led to local changes in existing LADs (Fig. 4e). Analysis of chromatin conformation and histone modifications showed that most of the gene promoters that resided in LADs were associated with increased open chromatin in K117fs iPSC-CMs compared to control iPSC-CMs (Fig. 4g and Extended Data Fig. 6f, g). Normalized assay for transposase accessible chromatin with high-throughput sequencing (ATAC-seq) and histone modification enrichment of each LAD were negatively correlated with enrichment by lamin A/C, and genes associated with LADs were more actively expressed in K117fs iPSC-CMs compared to control iPSC-CMs (Fig. 4h–k and Extended Data Fig. 6h–k). Moreover, histone H3 lysine-9 dimethylation (H3K9me2) in mutant iPSC-CMs was not equally distributed throughout the nuclear periphery and was less enriched in LADs region[28] (Extended Data Fig. 7a–e). Collectively, these data indicate that lamin A/C haploinsufficiency causes local changes in LADs leading to transcriptional activation.

Notably, we also found that many of the genes located in the non-LAD region were highly upregulated in K117fs compared to control iPSC-CMs, although the distance to the nearest LAD did not affect their expression (Extended Data Fig. 8a, b). The co-occurrence of transcription factors and ARCHS4 database analysis of differentially expressed genes located in non-LADs suggest that a prominent transcription factor of interest (*PRRX1*) that is located within a LAD may affect the abnormal expression of genes in non-LADs (Extended Data Fig. 8c–f). These data suggest a potential mechanism through which the alteration of LADs in K117fs iPSC-CMs affects the transcriptional regulation of genes located in non-LADs.

The PDGF pathway links to arrhythmic phenotype

To identify additional potential target genes that are closely associated with the disease phenotype, we compared the transcriptomes of K117fs mutant and control iPSC-CMs. By comparing the total RNA expression of control iPSC-CMs versus K117fs iPSC-CMs, we found that most of the differentially expressed genes were upregulated in K117fs iPSC-CMs (III-3, 84.87%; IV-1, 70.80%) (Fig. 5a). A cross-analysis of differentially expressed genes based on two different genetic backgrounds (III-3 and IV-1) identified 257 genes for which the expression in K117fs iPSC-CMs significantly differed from that in isogenic control iPSC-CMs (Fig. 5b). As expected, 239 out of 257 genes (93%) were upregulated in K117fs iPSC-CMs compared to isogenic control iPSC-CMs (Fig. 5c). Gene ontology (GO) enrichment analysis revealed that the upregulated genes in K117fs iPSC-CMs were functionally enriched in terms associated with platelet-derived growth factor (PDGF) binding arylsulfatase activity, protein binding involved in cell–matrix adhesion and PDGF receptor binding (Fig. 5d). The ARCHS4 kinase[29] analysis also showed that the upregulated genes in K117fs iPSC-CMs were highly enriched in the PDGF pathway.

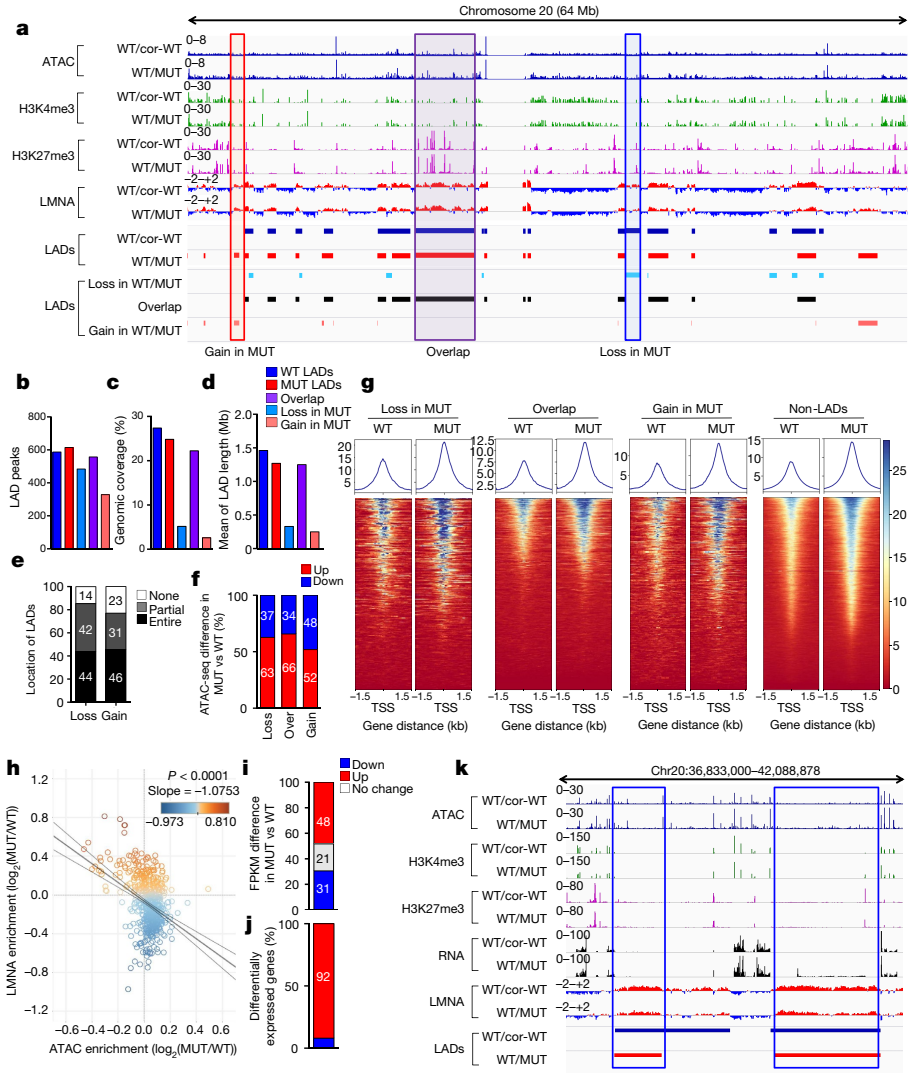


Fig. 4 | Lamin A/C haploinsufficiency results in reduced lamin A/C enrichment and increased open chromatin formation of each LAD. **a**, Representative images of chromatin immunoprecipitation followed by sequencing (ChIP–seq), ATAC-seq and RNA-seq of chromosome 20. The sc-376248 anti-LMNA antibody was used for ChIP–seq. **b**–**d**, Number (**b**), genomic coverage (**c**) and mean of length of LADs (**d**) in control, mutant, gained, overlapping and lost LADs. **e**, Location of LADs in the loss or gain category. LADs located within ±100 kb of overlapping LADs are shown as 'entire'. LADs partially shared with ±100 kb of overlapping LADs are showed as 'partial'. LADs located outside of ±100 kb of overlapping LADs are shown as 'none'. **f**, Comparison of normalized ATAC enrichment of each LAD in control and mutant iPSC-CMs. Red, percentage of LADs that showed upregulated normalized ATAC enrichment in mutant iPSC-CMs compared to control iPSC-CMs. Blue, percentage of LADs that showed downregulated normalized ATAC enrichment in mutant iPSC-CMs compared to control iPSC-CMs. **g**, Normalized ATAC-seq signal intensity around the transcription start site (TSS) of genes located in each LAD category. **h**, Scatter plot of normalized lamin A/C and ATAC enrichment of each LAD ($n = 588$). The y axis shows the $\log_2$-transformed relative normalized lamin A/C enrichment of each LAD in mutant iPSC-CMs compared to control iPSC-CMs. Note, in the graph, $\log_2$(MUT/WT) indicates $\log_2$(lamin A/C enrichment of each LAD in MUT/lamin A/C enrichment of each LAD in WT). The x axis shows the $\log_2$-transformed relative normalized ATAC enrichment of each LAD in mutant iPSC-CMs compared to control iPSC-CMs (shown as $\log_2$(MUT/WT)). Note, in the graph, $\log_2$(MUT/WT) indicates $\log_2$(ATAC enrichment of each LAD in MUT/ATAC enrichment of each LAD in WT). One dot represents one LAD. **i**, Comparison of normalized fragments per kb per million aligned reads (FPKM) of each LAD in control and mutant iPSC-CMs. Red, percentage of LADs with upregulated normalized FPKM in mutant iPSC-CMs compared to control iPSC-CMs; blue, percentage of LADs with downregulated normalized FPKM in mutant iPSC-CMs compared to control iPSC-CMs; grey, no change. **j**, Percentage of differentially expressed genes located in LADs. False-discovery rate (FDR)-corrected $P < 0.01$; $\log_2$-transformed fold change in expression of >1 or <-1. **k**, Representative images of ChIP–seq, ATAC-seq and RNA-seq. Blue boxes, LADs with lower enrichment of lamin A/C and higher expression in mutant iPSC-CMs compared to control iPSC-CMs.

The PDGF signalling is highly activated in smooth muscle and endothelial cells, and is initiated through the activation of two major receptors belonging to the PDGF receptor-α (PDGFRA) and PDGF receptor-β (PDGFRB) family[30]. During cardiomyocyte differentiation, *PDGFRA* and *PDGFRB* are highly upregulated in the early stages of differentiation but become downregulated after generating functional cardiomyocytes[31] (Extended Data Fig. 9a). In particular, expression of *PDGFRB* mRNA and PDGFRB protein is low in adult iPSC-CMs and normal heart tissues, but can be increased by stress conditions[32,33], which suggests that the PDGF signalling pathway is silenced in cardiomyocytes under physiological conditions (Extended Data Fig. 9a–c). However, we found that a significant increase in *PDGFRB* mRNA and protein expression occurred in K117fs iPSC-CMs compared to control iPSC-CMs (Fig. 5e–g and Extended Data Fig. 9d–f). In addition, a kinase array showed hyperactivation of PDGFRB in K117fs iPSC-CMs compared to isogenic control iPSC-CMs (Extended Data Fig. 9g). Furthermore, we found that the promoter region of the *PDGFRB* was more accessible in K117fs iPSC-CMs, as demonstrated by high enrichment of an active histone marker (H3K4me3) and open chromatin in

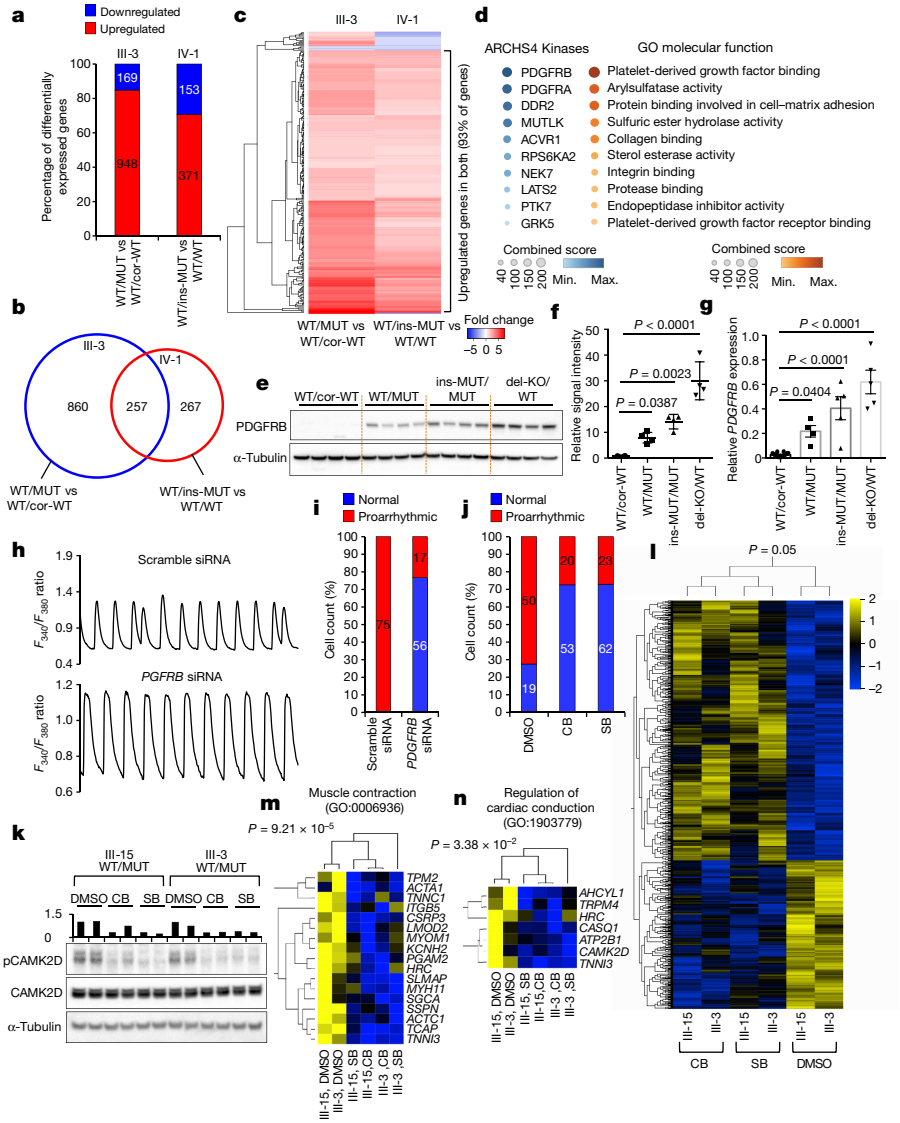


Fig. 5 | Abnormal activation of PDGFRB is required for the arrhythmic phenotype of mutant iPSC-CMs. a, Number of differentially expressed genes in mutant iPSC-CMs compared to control iPSC-CMs. LMNA WT/MUT and LMNA WT/cor-WT were derived from patient III-3. LMNA WT/WT and WT/ins-MUT were generated form health control IV-1. **b**, Venn diagram of differentially expressed genes in mutant iPSC-CMs compared to control iPSC-CMs. **c**, Heat maps of $\log_2$-transformed fold change in expression of 257 differentially expressed genes in mutant iPSC-CMs compared to control iPSC-CMs. **d**, GO and ARCHS4 kinase coexpression analysis of differential expressed genes. Colour codes indicate the combined FDR and *Z*-score. **e**, Immunoblot analysis of PDGFRB in control and mutant iPSC-CMs. **f**, Quantification of signal intensity of LMNA in **e**. $n = 4$. **g**, qPCR analysis of *PDGFRB* expression levels in control and mutant iPSC-CMs. $n = 8$ (WT/cor-WT), $n = 4$ (WT/MUT), $n = 5$ (ins-MUT/MUT and del-KO/MUT). **h**, Representative Ca^{2+} transients of mutant iPSC-CMs treated with scramble siRNA or siRNA against *PDGFRB*. All traces were recorded for 20 s. **i**, Quantification of cells that exhibit arrhythmic waveforms as shown in **h**. **j**, Quantification of cells that exhibit arrhythmic waveforms of Ca^{2+} transients for mutant iPSC-CMs treated with the PDGRB inhibitors crenolanib (CB) (100 nM) and sunitinib (SB) (500 nM) for 24 h. **k**, Immunoblot analysis of pCAMK2D and CAMK2D protein levels after treatment with dimethyl sulfoxide (DMSO), crenolanib or sunitinib. The analyses were independently repeated twice with similar results. **l**, Hierarchical clustering of amplicon-based sequencing (AmpliSeq) transcriptome data; analysed by one-way ANOVA ($P = 0.05$). Two different K117fs iPSC-CMs lines treated with crenolanib, sunitinib or DMSO were analysed by RNA-seq. The total number of genes is 915. **m**, **n**, GO analysis identified a set of genes that was related with muscle contraction and regulation of cardiac conduction. **f**, **g**, Data are mean ± s.e.m.; statistical significance was calculated using one-way ANOVA. The Ca^{2+} transients shown in **h** were independently repeated as described in **i** with similar results.

the ATAC-seq analysis (Extended Data Fig. 9h, i). Consistent with our observations in iPSC-CMs, heart tissue samples from both patients with LMNA-related DCM showed lower *LMNA* expression and higher *PDGFRB* expression when compared to healthy control tissues (Extended Data Fig. 9j, k). Taken together, these data suggest that PDGFRB is epigenetically activated in K117fs iPSC-CMs.

Next, we tested whether the abnormal activation of PDGFRB was directly linked to the arrhythmic phenotype that was observed in K117fs iPSC-CMs. Knockdown of *PDGFRB* expression in K117fs iPSC-CMs by small interfering (si)RNA resulted in a reduced prevalence of abnormal Ca^{2+} transients (23.28%, $n = 72$) compared to the treatment with scramble siRNA control (100%, $n = 75$) (Fig. 5h, i and Extended Data Fig. 10a–c). Treatment with two specific PDGFRB inhibitors, crenolanib and sunitinib, also ameliorated the arrhythmic phenotype of K117fs iPSC-CMs (crenolanib 27.39%, $n = 73$; sunitinib 27.05%, $n = 85$) compared to DMSO-treated cells (72.46%, $n = 69$) (Fig. 5j and Extended Data Fig. 10d–f). As expected, the phosphorylation of both CAMK2D and RYR2 was reduced after treatment of K117fs iPSC-CMs with crenolanib or sunitinib (III-15 and III-3) (Fig. 5k and Extended Data Fig. 10g). We also observed that the overexpression of PDGFRB resulted in upregulation of CAMK2D phosphorylation, inducing an arrhythmic phenotype in control iPSC-CMs (44.44%,

$n = 90$) (Extended Data Fig. 10h–j). These data indicate that the abnormal activation of PDGFRB contributes to the arrhythmic phenotype observed in K117fs iPSC-CMs.

To test the effects of the abnormal activation of PDGFRB on the gene-expression profile of K117fs iPSC-CMs, we next evaluated how treatment with crenolanib and sunitinib affected the transcriptome of K117fs iPSC-CMs. We identified a total of 910 genes that were differentially expressed between the treated and the untreated groups (Fig. 5l). GO term analysis of downregulated genes in the treated groups showed a high enrichment of genes related to heart functions, including muscle contraction, the regulation of cardiac conduction and ion transport (Fig. 5m, n and Extended Data Fig. 11a, b). We confirmed significant changes in the expression of genes related to cardiac muscle contraction and actin-mediated cell contraction through the knockdown of *PDGFRB* in K117fs iPSC-CMs (Extended Data Fig. 11c–e). We found that there were no differences in the lamin A/C level or the nuclear structure after treatment with crenolanib or sunitinib (Extended Data Fig. 11f–h). Taken together, these data confirm that the lamin A/C haploinsufficiency causes the abnormal activation of the PDGF signalling pathway, leading to the development of arrhythmias in LMNA-related DCM.

Discussion

Lamin A/C proteins are key components of heterochromatin conformation and the gene-silencing machinery, and are expressed in a cell-type-specific manner[23,34,35]. Here we elucidate how lamin A/C haploinsufficiency affects chromatin conformation and the gene-expression profile of *LMNA*-mutant iPSC-CMs. Furthermore, we demonstrate that the inhibition of the PDGF pathway ameliorates the arrhythmic phenotype of K117fs iPSC-CMs, suggesting a novel therapeutic target for the treatment of LMNA-related DCM (Extended Data Fig. 12). Our study suggests that several FDA-approved PDGFRB inhibitors—such as sunitinib, sorafinib and axitinib—may be repurposed for this condition. However, our previous study using a human iPSC-CM platform also revealed dose-dependent cardiac toxicity that is implicated in most tyrosine kinase inhibitors[36]. Therefore, further studies are warranted to identify the proper dosage or alternatives to these inhibitors that can be safely used in vivo to optimally alter the PDGF signalling pathway and prevent the fatal arrhythmias that are frequently observed in patients with LMNA-related DCM.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1406-x.

Received: 13 July 2017; Accepted: 19 June 2019;
Published online 17 July 2019.

1. Carmosino, M. et al. Role of nuclear lamin A/C in cardiomyocyte functions. *Biol. Cell* **106**, 346–358 (2014).
2. Fatkin, D. et al. Missense mutations in the rod domain of the lamin A/C gene as causes of dilated cardiomyopathy and conduction-system disease. *N. Engl. J. Med.* **341**, 1715–1724 (1999).
3. Krohne, G. & Benavente, R. The nuclear lamins. *Exp. Cell Res.* **162**, 1–10 (1986).
4. Hershberger, R. E. & Morales, A. in *GeneReviews* (eds Pagon, R. A. et al.) (University of Washington, 1993).
5. Hershberger, R. E., Hedges, D. J. & Morales, A. Dilated cardiomyopathy: the complexity of a diverse genetic architecture. *Nat. Rev. Cardiol.* **10**, 531–547 (2013).
6. Tesson, F. et al. Lamin A/C mutations in dilated cardiomyopathy. *Cardiol. J.* **21**, 331–342 (2014).
7. Diecke, S. et al. Novel codon-optimized mini-intronic plasmid for efficient, inexpensive, and xeno-free induction of pluripotency. *Sci. Rep.* **5**, 8081 (2015).
8. Kodo, K. et al. iPSC-derived cardiomyocytes reveal abnormal TGF-β signalling in left ventricular non-compaction cardiomyopathy. *Nat. Cell Biol.* **18**, 1031–1042 (2016).
9. Lee, J. et al. SETD7 drives cardiac lineage commitment through stage-specific transcriptional activation. *Cell Stem Cell* **22**, 428–444 (2018).
10. Burridge, P. W. et al. Chemically defined generation of human cardiomyocytes. *Nat. Methods* **11**, 855–860 (2014).
11. Karakikes, I. et al. A comprehensive TALEN-based knockout library for generating human induced pluripotent stem cell-based models for cardiovascular diseases. *Circ. Res.* **120**, 1561–1571 (2017).
12. Termglinchan, V., Seeger, T., Chen, C., Wu, J. C. & Karakikes, I. in *Cardiac Gene Therapy* (ed. Ishikawa, K.) 55–68 (Springer New York, 2017).
13. Bers, D. M. Calcium cycling and signaling in cardiac myocytes. *Annu. Rev. Physiol.* **70**, 23–49 (2008).
14. Lan, F. et al. Abnormal calcium handling properties underlie familial hypertrophic cardiomyopathy pathology in patient-specific induced pluripotent stem cells. *Cell Stem Cell* **12**, 101–113 (2013).
15. Itzhaki, I. et al. Modeling of catecholaminergic polymorphic ventricular tachycardia with patient-specific human-induced pluripotent stem cells. *J. Am. Coll. Cardiol.* **60**, 990–1000 (2012).
16. Maizels, L. et al. Patient-specific drug screening using a human induced pluripotent stem cell model of catecholaminergic polymorphic ventricular tachycardia type 2. *Circ Arrhythm Electrophysiol* **10**, e004725 (2017).
17. Bers, D. M. Cardiac sarcoplasmic reticulum calcium leak: basis and roles in cardiac dysfunction. *Annu. Rev. Physiol.* **76**, 107–127 (2014).
18. Schreiber, K. H. & Kennedy, B. K. When lamins go bad: nuclear structure and disease. *Cell* **152**, 1365–1375 (2013).
19. Kervestin, S. & Jacobson, A. NMD: a multifaceted response to premature translational termination. *Nat. Rev. Mol. Cell Biol.* **13**, 700–712 (2012).
20. Seeger, T. et al. A premature termination codon mutation in MYBPC3 causes hypertrophic cardiomyopathy via chronic activation of nonsense-mediated decay. *Circulation* **139**, 799–811 (2019).
21. Luperchio, T. R., Wong, X. & Reddy, K. L. Genome regulation at the peripheral zone: lamina associated domains in development and disease. *Curr. Opin. Genet. Dev.* **25**, 50–61 (2014).
22. Guelen, L. et al. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* **453**, 948–951 (2008).
23. Perovanovic, J. et al. Laminopathies disrupt epigenomic developmental programs and cell fate. *Sci. Transl. Med.* **8**, 335ra58 (2016).
24. Kind, J. & van Steensel, B. Genome–nuclear lamina interactions and gene regulation. *Curr. Opin. Cell Biol.* **22**, 320–325 (2010).
25. Chen, X. et al. ATAC-see reveals the accessible genome by transposase-mediated imaging and sequencing. *Nat. Methods* **13**, 1013–1020 (2016).
26. Gesson, K. et al. A-type lamins bind both hetero- and euchromatin, the latter being regulated by lamina-associated polypeptide 2 alpha. *Genome Res.* **26**, 462–473 (2016).
27. Rønningen, T. et al. Prepatterning of differentiation-driven nuclear lamin A/C-associated chromatin domains by GlcNAcylated histone H2B. *Genome Res.* **25**, 1825–1835 (2015).
28. Poleshko, A. et al. Genome–nuclear lamina interactions regulate cardiac stem cell lineage restriction. *Cell* **171**, 573–587 (2017).
29. Lachmann, A. et al. Massive mining of publicly available RNA-seq data from human and mouse. *Nat. Commun.* **9**, 1366 (2018).
30. Andrae, J., Gallini, R. & Betsholtz, C. Role of platelet-derived growth factors in physiology and medicine. *Genes Dev.* **22**, 1276–1312 (2008).
31. Tompkins, J. D. et al. Mapping human pluripotent-to-cardiomyocyte differentiation: methylomes, transcriptomes, and exon DNA methylation "memories". *EBioMedicine* **4**, 74–85 (2016).
32. Uhlén, M. et al. Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).
33. Chintalgattu, V. et al. Cardiomyocyte PDGFR-β signaling is an essential component of the mouse cardiac response to load-induced stress. *J. Clin. Invest.* **120**, 472–484 (2010).
34. Mattout, A., Cabianca, D. S. & Gasser, S. M. Chromatin states and nuclear organization in development — a view from the nuclear lamina. *Genome Biol.* **16**, 174 (2015).
35. Solovei, I. et al. LBR and lamin A/C sequentially tether peripheral heterochromatin and inversely regulate differentiation. *Cell* **152**, 584–598 (2013).
36. Sharma, A. et al. High-throughput screening of tyrosine kinase inhibitor cardiotoxicity with human induced pluripotent stem cells. *Sci. Transl. Med.* **9**, eaaf2584 (2017).

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators who performed electrophysiological tests and Ca^{2+} imaging analysis were blinded to group allocation during experiments and data collection. The studies comply with all ethical regulations.

Patient recruitment. The fibroblasts, PBMCs and heart tissues were obtained from patients using IRB-approved protocols at Stanford University (protocols 17576 and 29904). Informed consent was obtained from all patients who were included in our study. Clinical features of patients are described in the Extended Data Fig. 1d.

Culture and maintenance of iPSCs. iPSC lines were maintained in a chemically defined Essential 8 (E8 medium) medium (Life Technologies) on Matrigel-coated (BD Bioscience) plates at 37 °C with 5% (v/v) CO_2.

Pluripotency marker analysis. Human iPSC colonies grown in Matrigel-coated 8-well chamber glasses (Thermo Scientific) were fixed using 4% paraformaldehyde and permeabilized with 0.5% Triton X-100. After blocking samples with 5% goat serum in PBST (PBS with 0.1% Tween-20), cells were stained with mouse anti-SSEA4 (R&D systems), rabbit anti-OCT3/4 (Santa Cruz Biotechnology), rabbit anti-NANOG (Santa Cruz Biotechnology) and mouse anti-SOX2 (R&D systems) antibodies. Cells were then incubated with Alexa Fluor-conjugated secondary antibodies (Life Technologies) and Hoechst 33342 (Life Technologies) to visualize the specific stains. Image acquisition was performed on an Eclipse 80i fluorescence microscope (Nikon Instruments).

TALEN-mediated homologous recombination. TALEN pair vectors were designed and constructed using the rapid TALEN assembly system as previously described[11]. In brief, 500 base-pair (bp) fragments of wild-type *LMNA* exon 1 and adjacent intronic sequences were synthesized as GeneArt String DNA fragments (Life Technologies) to make left and right homologous arms, and cloned into PB-MV1Puro-TK vectors (Transposagen), as previously described[12]. Two silent mutations in the homologous arms were inserted to avoid recleavage of the genomic sequence. Both TALEN pairs and targeting vectors were delivered into iPSCs by nucleofection using P3 Primary Cell 4D-Nucleofector X Kit (Lonza). Afterwards, cells with the correct targeting vector integration were selected by puromycin (Life Technologies) and genotyped. To excise the selection cassette, transient expression of piggyBac transposase was performed by transfection of excising piggyBac transposase mRNA (Transposagen) using Lipofectamine MessengerMAX (LifeTechnologies). After negative selection using ganciclovir (Sigma Aldrich), the established clones were genotyped by PCR and bidirectional direct sequencing.

TALEN-mediated non-homologous end joining. TALEN pair vectors were designed and constructed using the rapid TALEN assembly system as previously described[11,12] and were delivered into iPSCs by nucleofection using the P3 Primary Cell 4D-Nucleofector X Kit (Lonza). Subsequently, 45 h after nucleofection, transfected cells were enriched by fluorescence-activated cell sorting (FACS) and established clones were genotyped by PCR and bidirectional direct sequencing.

Off-target detection. Genomic DNA was extracted from gene-edited iPSC clones using the DNeasy Blood & Tissue Kit (Qiagen). The potential TALEN off-target sites were predicted in silico based on sequence homology using the bioinformatics tool PROGNOS. The top 20 targets were investigated by DNA sequencing. The primers designed by PROGNOS were used to amplify the genomic regions of putative off-target sites by PCR. Each PCR reaction contained 1.25 units of Prime STARGXL DNA Polymerase (Clontech) and 50 ng of genomic DNA (total volume 20 μl). The PCR products were analysed by Sanger sequencing and sequencing reads were aligned to the wild-type sequence obtained from the parental iPSC line.

Immunocytochemistry. Cells grown on coverslips were fixed using 4% paraformaldehyde, permeabilized with 0.5% Triton X-100, incubated with primary antibodies and Hoechst 33342, and detected using Alexa Fluor-conjugated secondary antibodies. Primary antibodies include rabbit anti-cardiac troponin T (Abcam), mouse anti-cardiac troponin T (Thermo Scientific), mouse anti-sarcomeric α-actinin (Sigma-Aldrich), goat anti-LMNA (Santa Cruz) and rabbit anti-LMNA (Santa Cruz) antibodies. Image acquisition was performed on an Eclipse 80i fluorescence microscope, a confocal microscope (Carl Zeiss, LSM 510 Meta) and ZEN software (Carl Zeiss).

Reverse transcription and quantitative PCR. Total mRNA was isolated from iPSC-CMs using the Qiagen miRNeasy Mini kit. Subsequently, 1 μg of RNA was used to synthesize cDNA using the iScript cDNA Synthesis kit (Bio-Rad). Then, 0.25 μl of the reaction was used to quantify gene expression by qPCR using TaqMan Universal PCR Master Mix. Expression values were normalized to the average expression of the housekeeping gene 18S.

Western blotting. Proteins were resolved by SDS–PAGE and were transferred to 0.45-μm nitrocellulose membranes (Bio-Rad) using a mini Bio-Rad Mini PROTEAN 3 Cell system in NuPAGE transfer buffer (Life Technologies). The membrane was then blocked in Membrane Blocking Solution (Life Technologies) and incubated with primary antibodies overnight at 4 °C. Blots were incubated with the appropriate secondary antibodies for 1 h at room temperature and visualized using the ECL Western Blotting Analysis System (GE Healthcare). Primary antibodies used were mouse anti-LMNA (Santa Cruz), rabbit anti-LMNA (Santa Cruz), CAMK2D (Abcam), PDGFRB (Cell Signaling), RYR2 (Abcam), pRYR2 (D. M. Bers laboratory) and HRP-conjugated α-tubulin (Cell Signaling).

Patch-clamp recordings. Whole-cell action potentials were recorded using a standard patch-clamp technique. In brief, cultured iPSC-CMs were plated on No. 1 18-mm glass coverslips (Warner Instruments) coated with Matrigel, placed in a RC-26C recording chamber (Warner Instruments) and mounted onto the stage of an inverted microscope (Nikon). The chamber was continuously perfused with warm (35–37 °C) extracellular solution (pH 7.4) of the following composition: NaCl (140 mM), KCl (5.4 mM), $CaCl_2$ (1.8 mM), $MgCl_2$ (1 mM), HEPES (10 mM) and glucose (10 mM); pH was adjusted to 7.4 with NaOH. Glass micropipettes were fabricated from standard wall borosilicate glass capillary tubes (Sutter BF 100-50-10, Sutter Instruments) using a programmable puller (P-97; Sutter Instruments) and filled with the following intracellular solution (in mM): 120 KCl, 1.0 $MgCl_2$, 10 HEPES, 10 EGTA and 3 Mg-ATP (pH 7.2). A single beating cardiomyocyte was selected and action potentials were recorded in whole-cell current-clamp mode using an EPC-10 patch-clamp amplifier (HEKA). Data were acquired using Patch Master software (HEKA) and digitized at 1.0 kHz.

Differentiation of iPSC-CMs. iPSCs were grown to 90% confluence and subsequently differentiated into beating cardiomyocytes, using a small-molecule-based monolayer method that has previously been described[10]. After ten days of cardiac differentiation, iPSC-CM monolayers were purified using RPMI-1640 without glucose (Life Technologies) and with B-27 supplement (Life Technologies). The non-glucose culture medium was changed every two days. After five days, iPSC-CMs were reseeded on Matrigel-coated plates in a culture medium containing glucose.

siRNA-mediated knockdown. Gene knockdown experiments were performed using Lipofectamine RNAiMax (Life Technologies) according to the manufacturer's instructions. Cells were transfected with either scramble siRNA or siRNA against *PDGFRB* (SilencerRSelect, ThermoFisher, 25 nM per well, 4390824) for 48 h before being subjected to subsequent downstream analyses.

Treatment with NMD inhibitors. The potent NMD inhibitors emetine and wortmannin (Sigma-Aldrich) were dissolved in water and DMSO, respectively. An equal concentration of solvent (water or DMSO) was used as the control. iPSC-CMs were treated with emetine or wortmannin for 6 h before the experiment.

Treatment with PDGFRB inhibitors. The PDGFRB inhibitors sunitinib and CP-868596 (Selleckchem) were dissolved in DMSO. An equal concentration of solvent (DMSO) was used as the control. iPSC-CMs were treated with sunitinib or CP-868596 for 48 h before the experiment.

Treatment with CAMK2D inhibitors. The active CAMK2D inhibitor (KN93) and inactive CAMK2D inhibitor (KN92) were dissolved in DMSO. iPSC-CMs were treated with KN92 or KN93 for 24 h before the experiment.

Droplet digital PCR. Total RNA was extracted from iPSC-CMs at day 30 post-differentiation using the miRNeasy Mini Kit (QIAGEN) and cDNA preparation was carried out using the iScript cDNA Synthesis Kit (Bio-Rad Laboratories). The concentration of cDNA was reduced to about 0.2 ng μl^{-1} RNA equivalent, and 1 ng (5 μl of 0.2 ng μl^{-1}) of RNA-equivalent cDNA was mixed with primers, probes and ddPCR Supermix reaction (total volume 20 μl). The final concentrations of the primers and the probe were 900 nM and 500 nM, respectively. The following primers and probes for discriminating allelic expression of *LMNA* K117 (wild-type allele) from K117fs (mutant allele) were used: forward primer, 5′-GCAAGACCCTTGACTCAGTA-3′; reverse primer, 5′-CTCCTTGGAGTTCAGCAG-3′; wild-type probe: 5′(6-FAM)-TGCGCGCTTTCAGCTCCTTAA-(Blackhole Quencher)3′; and mutant probe, 5′(HEX)-TGCGCGCTTTCCAGCTCCT-(Blackhole Quencher)3′. Droplet formation was carried out using a QX100 droplet generator. A rubber gasket is placed over the cartridge and loaded into the droplet generator. The emulsion (35 μl in volume) was then slowly transferred using a multichannel pipette to a 96-Well twin.tec PCR Plate (Eppendorf). The plate was heat-sealed with foil and the emulsion was cycled to end point per the manufacturer's protocol with an annealing temperature at 61 °C. Finally, the samples were analysed using a BioRad QX100 reader.

Ca^{2+} imaging. iPSC-CMs seeded on a glass coverslip for 5–7 days were loaded with the cell-permeable calcium-sensitive dye fura-2 AM (2 $\mu mol\ l^{-1}$) for 20 min. After 15 min of washing in 1.8 mmol l^{-1} Ca^{2+}-Tyrode (135 mmol l^{-1} NaCl, 4 mmol l^{-1} KCl, 1 mmol l^{-1} $MgCl_2$, 5 mmol l^{-1} glucose and 10 mmol l^{-1} HEPES, pH 7.4) buffer to allow de-esterification, coverslips were mounted on the stage of an inverted epifluorescence microscope (Nikon Eclipse Ti-S). iPSC-CMs were field-stimulated at 0.5 Hz with a pulse duration of 10 ms. Fura-2-AM-loaded cells were excited at both 340 and 380 nm, and the emission fluorescence signal was collected at 510 nm as previously described[37]. Changes in fluorescence signal were measured using the NIS Elements AR software, which permits the recording of multiple cells in one view. Intracellular calcium changes were expressed as changes in the ratio $R = F_{340}/F_{380}$ and the calcium transient waves were analysed using a previously published method[38].

Measuring abnormal calcium release from the sarcoplasmic reticulum. Both patient- and healthy individual-derived iPSC-CMs were seeded on coverslips as single cells. After 3–4 days of recovery, the cells were loaded with 5 μM Fluo-4 AM at 37 °C for 10 min and then washed with Tyrode's solution three times. Ca^{2+} release events were recorded with a Carl Zeiss confocal (710) in line-scanning mode (512 pixels × 1,920 lines). The extracellular media were prepared with sequential increases of Ca^{2+} concentration (0, 0.5, 1, 2 and 5 mM), and were used to treat iPSC-CMs during the recording. The Ca^{2+} imaging data were displayed and analysed using Image J.

Measuring sarcomeric alignments. Immunostaining images of iPSC-CMs were viewed with Image J, and the fluorescent signals along the sarcomere structure were pulled out. A custom-made Interactive Digital Language algorithm was used to analyse the regularity of sarcomere signal distribution with fast Fourier transformation (FFT). The sarcomere length and the regularity of sarcomere distribution were indicated as the position and the height of the first main peak after FFT data processing.

ChIP–seq. LMNA antibodies (Santa Cruz Biotechnology; sc-376248 and Abcam; 8984) were incubated with Dynabeads (Life Technologies; 10003D) for 12 h at 4 °C. A small portion of the crosslinked, sheared chromatin was saved as the input, and the remainder was used for immunoprecipitation using antibody-conjugated Dynabeads. After overnight incubation at 4 °C, the incubated beads were rinsed with sonication buffer (50 mM HEPES pH 7.9, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS, 0.5 mM PMSF), a high-salt buffer (50 mM HEPES pH 7.9, 500 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS, 0.5 mM PMSF) and a LiCl buffer (20 mM Tris, pH 8.0, 1 mM EDTA, 250 mM LiCl, 0.5% NP-40, 0.5% sodium deoxycholate, 0.5 mM PMSF). The washed beads were incubated with elution buffer (50 mM Tris, pH 8.0, 1 mM EDTA, 1% SDS, 50 mM $NaHCO_3$) for 1 h at 65 °C and then de-crosslinked with 5 M NaCl overnight at 65 °C. The immunoprecipitated DNA was treated with RNase A and proteinase K and purified by ChIP DNA Clean and Concentrator (Zymo Research; D5205). The raw sequencing data were analysed as previously described[39].

RNA-seq. For each sample in the whole-transcriptome sequencing library, 60–80 million 75-bp paired-end reads were acquired from the sequencer. Base quality of raw reads is high after checking with FastQC 0.11.4. Using STAR 2.5.1b, we aligned the reads to the human reference genome (hg19), with splice junctions defined by the GTF file downloaded from UCSC. On average, 92% of reads were aligned to the reference genome, and 83% of reads were uniquely aligned to the reference genome. Gene expression was determined by calculating the FPKM using Cufflinks 2.2.1. In addition, Cufflinks was used to determine differential expression between each two conditions.

ATAC-seq. The samples were treated and processed as previously described[40]. In brief, 100,000 cells were centrifuged at 500*g* for 5 min at room temperature. The cell pellet was resuspended in 50 ml lysis buffer (10 mM Tris-Cl pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.01% Igepal CA-630) and centrifuged immediately at 500*g* for 10 min at 4 °C. The cell pellet was resuspended in 50 ml transposase mixture (25 μl 2× TD buffer, 22.5 μl, dH_2O and 2.5 μl Illumina Tn5 transposase or 100 nM (final concentration) Atto-590-labelled in-house-generated Tn5) and incubated at 37 °C for 30 min. After transposition, the mixture was purified with the Qiagen Mini purification kit and eluted in 10 μl Qiagen EB elution buffer. Sequencing libraries were prepared following the original ATAC-seq protocol[40]. The sequencing was performed on Illumina NextSeq at the Stanford Functional Genomics Facility. ATAC-seq reads were trimmed of adapters and then mapped to hg19 genome assembly using Bowtie 2[41]. Following quality control to remove duplicate reads, average read intensities were calculated with the aid of deepTools[42] and R/Bioconductor (v.3.2.1)[43]. Promoter regions were defined as ±1 kb around the hg19 gene transcription start site coordinates unless otherwise stated.

ATAC-see. The samples were treated and processed as previously described[25]. In brief, iPSC-CMs were fixed with 1% formaldehyde (Sigma) for 10 min and quenched with 0.125 M glycine for 5 min at room temperature. After fixation, the cells (either growing on slides or centrifuged on glass slides with Cytospin) were permeabilized with a lysis buffer (10 mM Tris-Cl pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.01% Igepal CA-630) for 10 min at room temperature. After permeabilization, the slides were rinsed in PBS twice and placed in a humid chamber box at 37 °C. The transposase mixture solution (25 μl 2× TD buffer, final concentration of 100 nM Tn5-ATTO-59ON, adding dH_2O up to 50 μl) was added to the slide and incubated for 30 min at 37 °C. After the transposase reaction, slides were washed with PBS containing 0.01% SDS and 50 mM EDTA for 15 min three times at 55 °C. After washing, slides were mounted using Vectashield with DAPI (H-1200, Vector Laboratories). Fluorescence images were captured on a confocal microscope equipped with a 40× oil-immersion lens. Fluorescent intensity profiles of DAPI and ATAC were exported using ZEN (Zeiss). To find out whether the *LMNA* mutation led to the specific re-distribution of the epigenetic histone markers in the iPSC-CMs, the correlation between ATAC-see and DAPI signal of each nucleus was calculated using the Pearson correlation method. Image analysis was conducted using Graphpad Prism v.7.0.

RNA-seq and ChIP–seq analysis. FastQC (v.0.11.5) and MultiQC (v.1.3) were used to assess read quality. Adaptor and quality trimming of reads were performed with trimmomatic (v.0.36). Reads were mapped to the hg19 reference genome using STAR (v.2.5.3a) with ENCODE long RNA-seq parameters. Uniquely mapped reads were filtered for and bigWig files were generated with samtools (v.1.4). FPKM values and differentially expressed genes were obtained using cuffdiff (v.2.2.1). ChIP–seq data were processed using the AQUAS pipeline from the Kundaje laboratory at Stanford University (https://github.com/kundajelab/chip-seq-pipeline2), which has an end-to-end implementation of the ENCODE (phase 3) ChIP–seq pipeline. Default parameters were used with the exception of specifying '-type histone -species hg19'. Before LAD detection of LMNA data, duplicate reads were removed with 'mark duplicates' from Picard tools (v.2.17.3) and 'DownsampleSam' was used to downsample the larger of each pair of aligned input and ChIP read files, giving each pair the same read depth and avoiding normalization bias.

LAD detection and analysis. Lamin A/C binding data were analysed using Enriched Domain Detector[39] (v.1.0) with an 11-kb bin size, gap penalty of 5, and FDR-adjusted significance threshold of $P < 0.05$. Gains, losses and intersections in LADs between control and mutant cells were tallied using bedtools (v.2.27.1). Gene-expression changes within each category of LADs (gain, loss or shared in mutant and control) based on RNA-seq data were compared in R v.3.2.1 and Bioconductor[43] using the iRanges and GenomicRanges packages[44]. In deciding whether a gene overlaps with each category, the union of called peaks from the lamin A/C ChIP–seq of two antibodies and the intersection of two cell lines were used. A gene is considered to reside in a particular LAD if any of its hg19-annotated transcription start sites overlaps with the LAD range by genomic coordinates. In cases of ambiguity, intersection (shared between control and mutant cells) regions take precedence over gain and loss regions. ATAC-seq read intensities within each LAD category around genomic features, including transcription start sites and transcription end sites, were visualized using deepTools[42] with feature coordinates from hg19 annotations.

Statistical analyses. Data were expressed as mean ± s.e.m. Immunoblots are representative of at least two independent experiments. All other experiments are the average of at least 2 independent assays, and for cell number calculations in immunostaining assays, at least 100 cells per sample were counted for each independent experiment. Statistical analyses were performed using GraphPad Prism (v.6.0e). An unpaired two-tailed Student's *t*-test was used to calculate significant differences between two groups. Multiple comparison correction analysis was performed using one-way ANOVA followed by Tukey's post hoc HSD test. $P < 0.05$ was considered statistically significant.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Data are available from the Gene Expression Omnibus (GEO; GSE118885).

37. Lam, C. K. et al. Novel role of HAX-1 in ischemic injury protection involvement of heat shock protein 90. *Circ. Res.* **112**, 79–89 (2013).
38. Greensmith, D. J. Ca analysis: an Excel based program for the analysis of intracellular calcium transients including multiple, simultaneous regression analysis. *Comput. Methods Programs Biomed.* **113**, 241–250 (2014).
39. Lund, E., Oldenburg, A. R. & Collas, P. Enriched domain detector: a program for detection of wide genomic enrichment domains robust against local variations. *Nucleic Acids Res.* **42**, e92 (2014).
40. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013).
41. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
42. Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42**, W187–W191 (2014).
43. Huber, W. et al. Orchestrating high-throughput genomic analysis with Bioconductor. *Nat. Methods* **12**, 115–121 (2015).
44. Lawrence, M. et al. Software for computing and annotating genomic ranges. *PLOS Comput. Biol.* **9**, e1003118 (2013).

Acknowledgements We thank members of the laboratory of D. M. Bers for providing the pRYR2 antibody, and S. A. Yi, K. H. Nam, G. A. Akgun, C. Chen and S. Zhang for their contribution. This research is supported by AHA 17MERIT33610009, NIH R01 HL128170, R01 HL113006, R01 HL130020, R01 HL132875, R01 HL141851, Leducq Foundation 18CVD05 (J.C.W.); R01 HL139679, R00 HL104002, AHA 17IRG33410532 (I.K.); Prince Mahidol Award Foundation (V.T.); NIH K99 HL133473 (H.W.); the German Research Foundation (T.S.); National Research Foundation of

Korea 2012R1A5A2A28671860 and 2019R1C1C1010675 (J.L.); NIH F32 HL139045 (E.L.); CIRM GC1R-06673-A, R24 HL117756 (M.P.S.); and NIH P50 HG007735 and Howard Hughes Medical Institute (H.Y.C.)

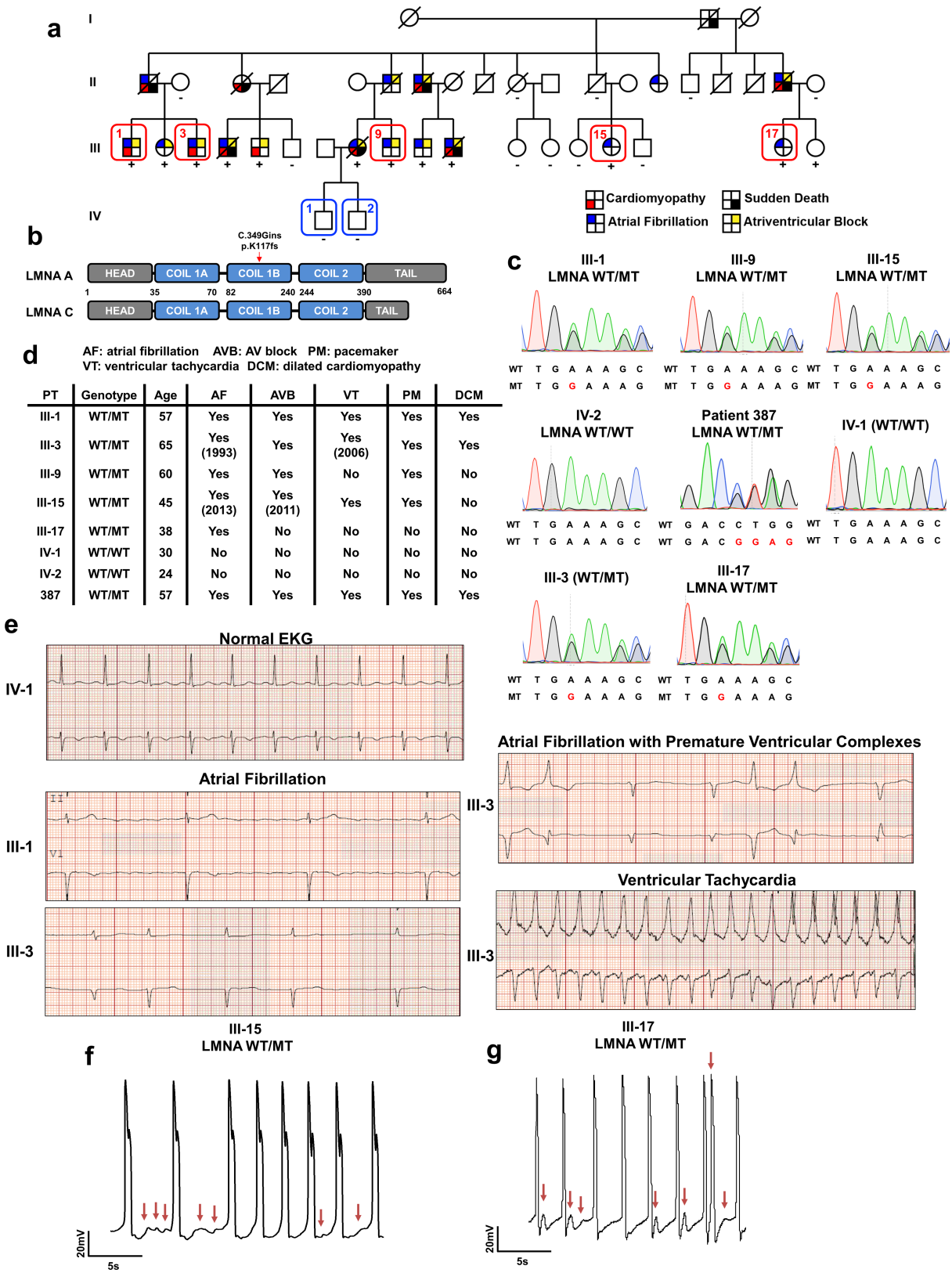
Author contributions J.C.W., I.K, V.T. and S.D. conceptualized and designed the study; J.C.W., I.K., J.L., V.T. and S.D. wrote the manuscript; V.T. and I.K. designed and performed the gene-editing experiments; J.L. analysed the majority of the data; I.I. and P.G. performed the electrophysiological experiments; C.K.L. and H.W. performed calcium analysis; X.C. and J.Z.Z. performed the ATAC-see experiments; M.A. performed the ATAC-seq experiments; M.G., E.L., J.L. and J.M.C. analysed all of the next-generation sequencing-related data; K.S. and J.-W.R. performed skin biopsies and clinical phenotype assessments; P.J.W. recruited the patients; J.L., R.C., T.S., T.C. and I.P.G. generated iPSC-CMs for the experiments; M.P.S. and H.Y.C. contributed to the experimental design; I.K. and J.C.W. supervised the study and provided funding support.

Competing interests H.Y.C. is a advisor to10x Genomics. Stanford University has filed a patent application on ATAC-see technology, in which H.Y.C. is named a co-inventor. J.C.W. is a co-founder of Khloris Biosciences but has no competing interests, as the work presented was performed independently.

Additional information

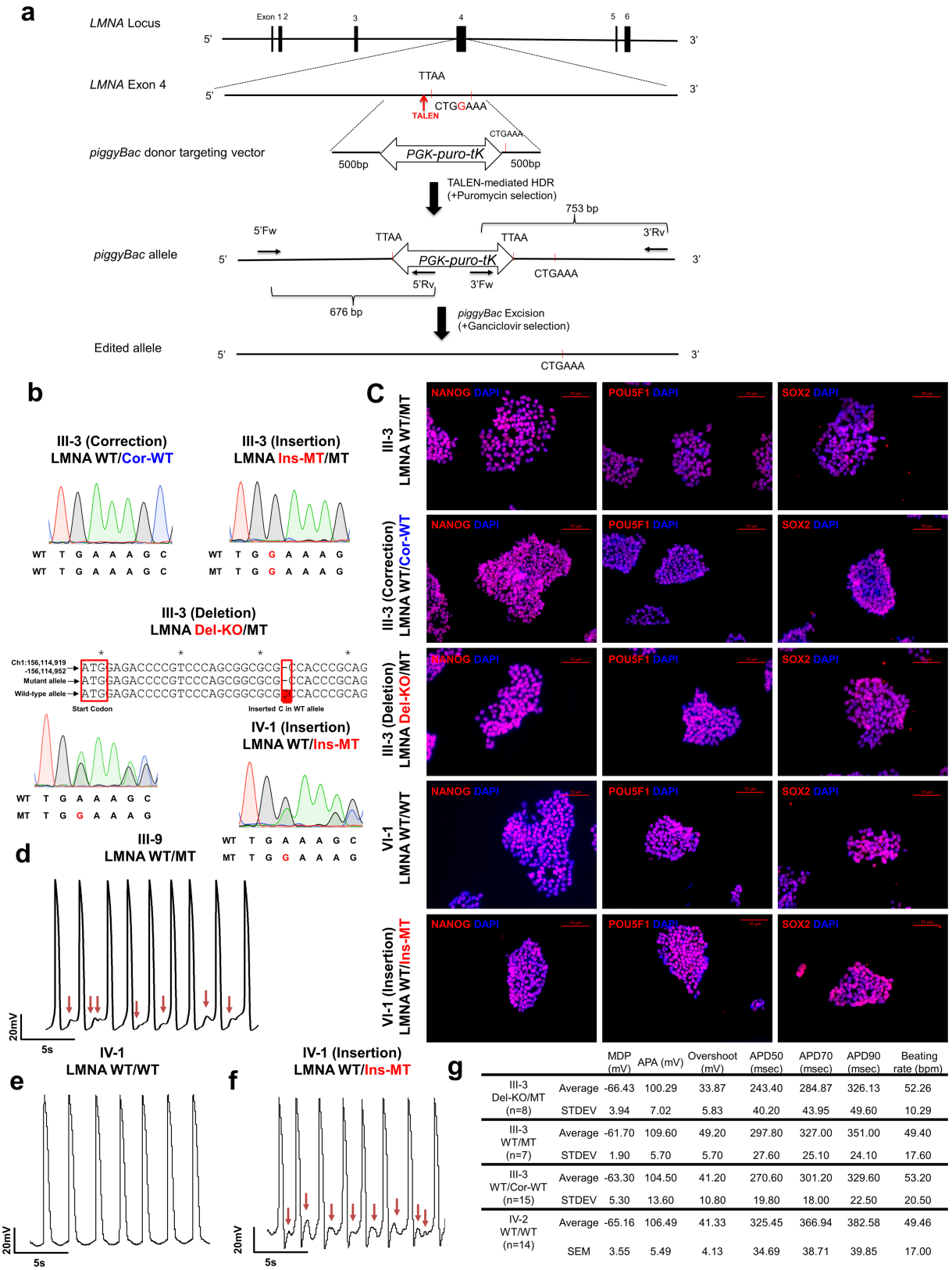
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1406-x.

Correspondence and requests for materials should be addressed to J.L., I.K. or J.C.W.



PT	Genotype	Age	AF	AVB	VT	PM	DCM
III-1	WT/MT	57	Yes	Yes	Yes	Yes	Yes
III-3	WT/MT	65	Yes (1993)	Yes	Yes (2006)	Yes	Yes
III-9	WT/MT	60	Yes	Yes	No	Yes	No
III-15	WT/MT	45	Yes (2013)	Yes (2011)	Yes	Yes	No
III-17	WT/MT	38	Yes	No	No	No	No
IV-1	WT/WT	30	No	No	No	No	No
IV-2	WT/WT	24	No	No	No	No	No
387	WT/MT	57	Yes	Yes	Yes	Yes	Yes

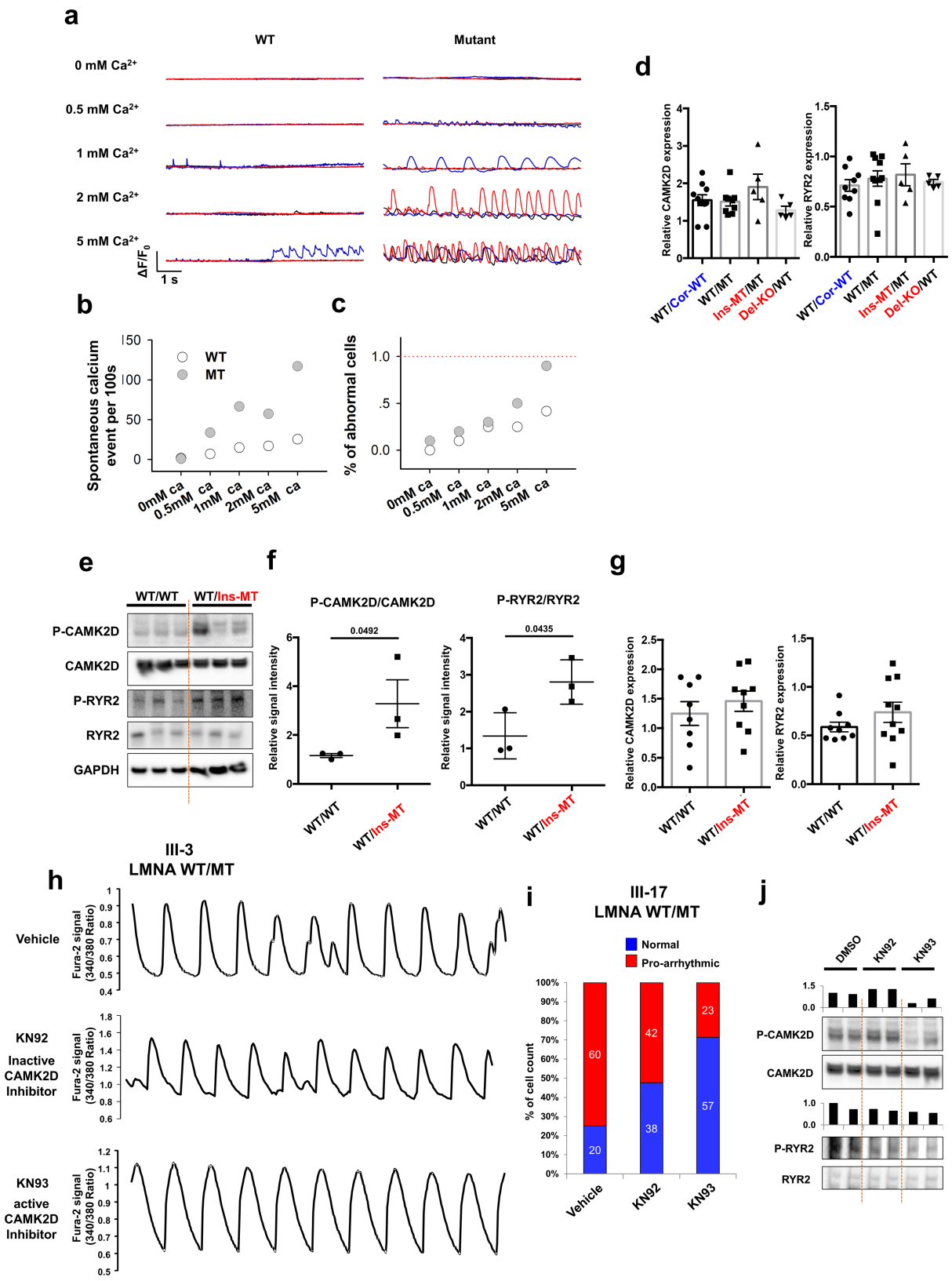
Extended Data Fig. 1 | *LMNA*-mutant iPSC-CMs can recapitulate arrhythmic phenotype of patients with LMNA-related DCM. **a**, Schematic pedigree of the family carrying the mutation in *LMNA*. Patients (III-1, III-3, III-9, III-15 and III-17) and healthy individuals (IV-1 and IV-2) recruited for this study are numbered. Circles represent female family members and squares represent males. The '+' and '−' signs underneath family members indicate the presence or absence of the mutation in *LMNA*, respectively. **b**, Schematic view of 348-349insG frameshift mutation in *LMNA*. **c**, Genotyping of fibroblasts derived from patients and healthy controls. MT, mutant. **d**, Clinical features of patients and healthy individuals. **e**, Electrocardiogram of patients III-1 and III-3 and a healthy individual (IV-1). The electrocardiogram data were measured once per individual. **f**, **g**, Electrophysiological measurements of spontaneous action potentials in mutant iPSC-CMs (III-15 and III-17) recorded by patch clamp in current-clamp mode. The experiments were repeated three times independently with similar results.



g

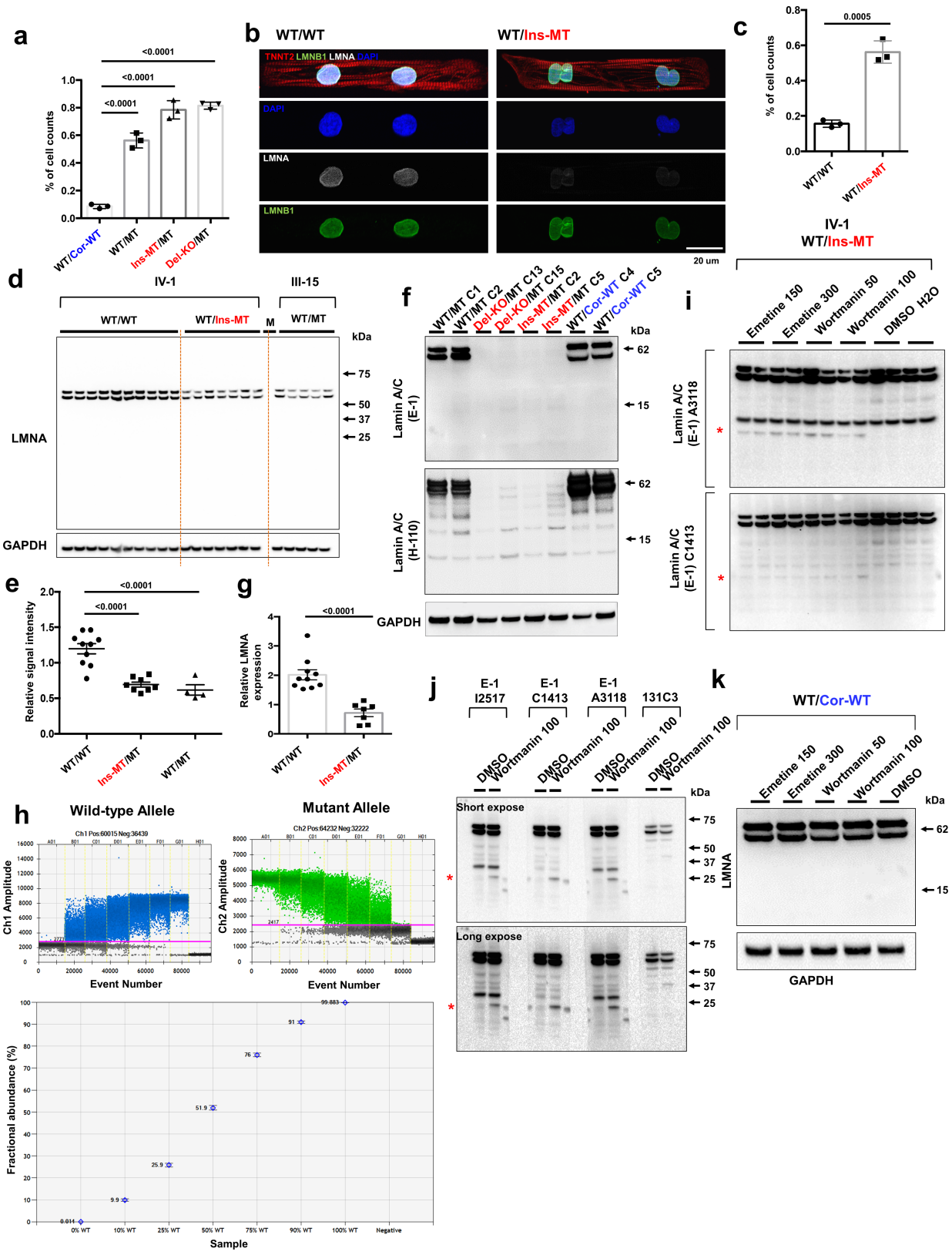
		MDP (mV)	APA (mV)	Overshoot (mV)	APD50 (msec)	APD70 (msec)	APD90 (msec)	Beating rate (bpm)
III-3 Del-KO/MT (n=8)	Average	-66.43	100.29	33.87	243.40	284.87	326.13	52.26
	STDEV	3.94	7.02	5.83	40.20	43.95	49.60	10.29
III-3 WT/MT (n=7)	Average	-61.70	109.60	49.20	297.80	327.00	351.00	49.40
	STDEV	1.90	5.70	5.70	27.60	25.10	24.10	17.60
III-3 WT/Cor-WT (n=15)	Average	-63.30	104.50	41.20	270.60	301.20	329.60	53.20
	STDEV	5.30	13.60	10.80	19.80	18.00	22.50	20.50
IV-2 WT/WT (n=14)	Average	-65.16	106.49	41.33	325.45	366.94	382.58	49.46
	SEM	3.55	5.49	4.13	34.69	38.71	39.85	17.00

Extended Data Fig. 2 | The mutation in *LMNA* is a cause of the arrhythmic phenotype in *LMNA*-mutant iPSC-CMs. a, Gene-editing strategy using the TALEN method. The piggyBac system was used to generate isogenic lines as previously described[11,12]. **b**, Genotyping of gene-edited isogenic lines (III-3 corrected, insertion, deletion; IV-1 insertion). For LMNA del-KO/MUT, we used TALEN pairs that target the start codon of *LMNA*. Genotyping showed the C insertion in the wild-type allele that leads early stop codon. **c**, Immunostaining of NANOG (red, left), POU5F1 (red, middle) and SOX2 (red, right) in iPSC lines. Blue, DAPI. Scale bars, 10 μm. The experiments were repeated twice independently with similar results. **d**–**f**, Electrophysiological recordings of spontaneous action potentials in control (IV-1) and mutant (III-9, isogenic IV-1; WT/ins-MUT) iPSC-CMs measured by patch clamp in current-clamp mode. Red arrows indicate delayed afterdepolarization-like arrhythmias. The experiments were repeated three times independently with similar results. **g**, Action potential parameters of ventricular-like iPSC-CMs. MDP, maximal diastolic potential; APA, action potential amplitude; APD, action potential duration at 50%, 70%, 90% of repolarization; bpm, beats per min.



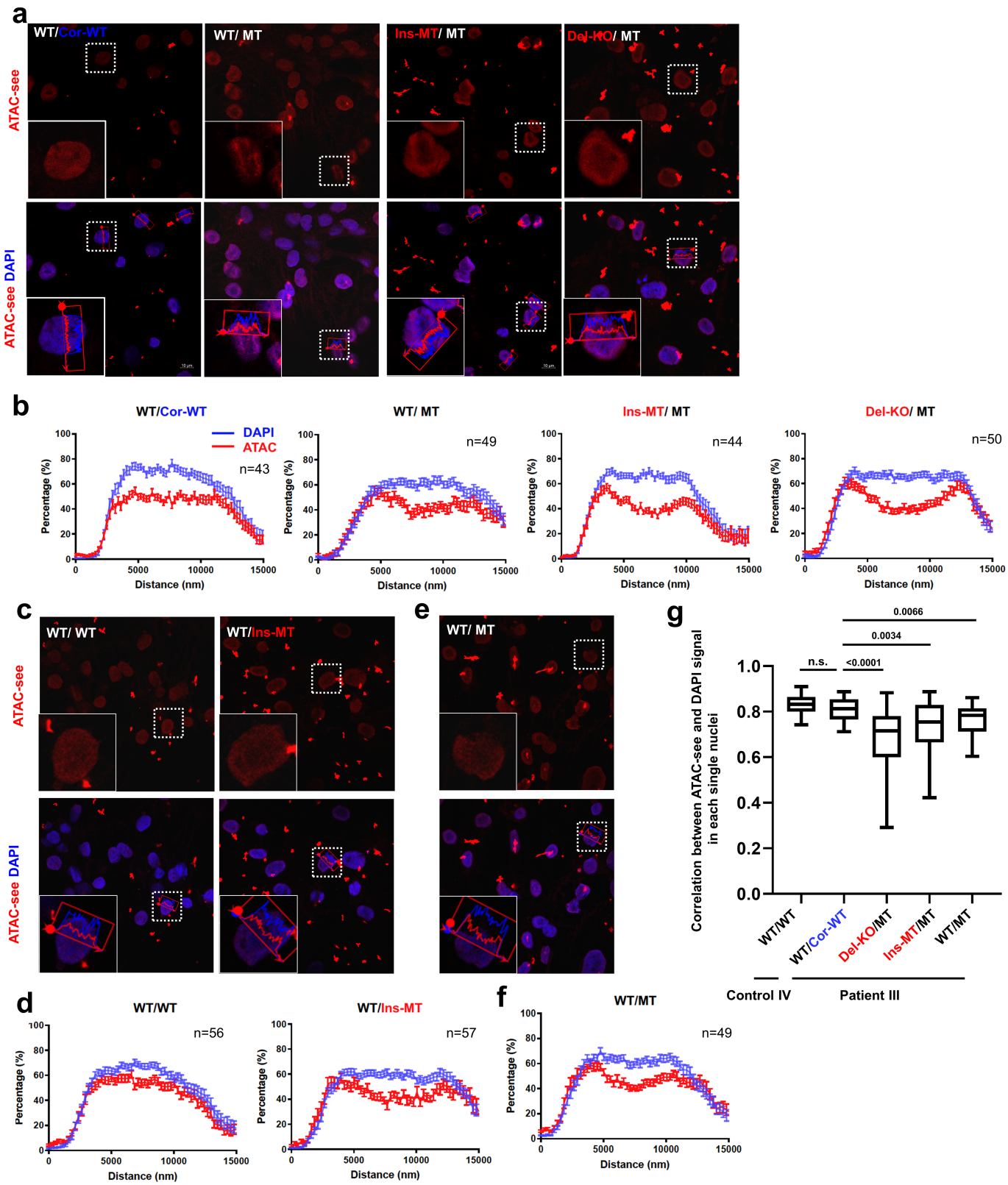
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Abnormal calcium handling in *LMNA*-mutant iPSC-CMs. a, Confocal imaging of Fluo-4 AM calcium events in control (III-3; WT/cor-WT) and mutant (III-3; WT/MUT) iPSC-CMs while being treated with increasing extracellular Ca^{2+} concentrations. All representative traces were recorded from three individual cells (presented as red, blue and black). **b**, Spontaneous calcium events per 100 s of control and mutant iPSC-CMs for each extracellular Ca^{2+} concentration. **c**, Summary of the percentage of cells that have spontaneous Ca^{2+} release events from the sarcoplasmic reticulum in control and mutant iPSC-CMs. **d**, qPCR analysis of *CAMK2D* and *RYR2* expression in control and mutant iPSC-CMs. Data are mean ± s.e.m. **e**, **f**, Immunoblot analysis of pRYR2, RYR2, pCAMK2D and CAMK2D protein levels in control and mutant iPSC-CMs. Data are mean ± s.e.m.; a two-tailed Student's *t*-test was used to calculate *P* values; $n = 3$; values above the lines indicate significance. **g**, qPCR analysis of *CAMK2D* expression in control and mutant iPSC-CMs. Expression level of *GAPDH* was used as control. Data are mean ± s.e.m.; $n = 8$. **h**, Representative Ca^{2+} transients of mutant iPSC-CMs (III-3; WT/MUT) treated with 1 μM of KN92 or KN93 for 24 h. **i**, Quantification of the percentage of cells that exhibit arrhythmic waveforms in mutant iPSC-CMs (III-17 WT/MUT) at baseline, as well as after the treatment with 1 μM of KN92 or KN93 for 24 h. **j**, Immunoblot analysis of pRYR2, RYR2, pCAMK2D and CAMK2D protein levels after treatment of DMSO, KN92 or KN93 for 24 h. The experiments in **a** were repeated twice independently with similar results. The Ca^{2+} transient analyses in **h** were repeated as described in Fig. 2e independently with similar results. The immunoblot analyses in **e** and **j** were repeated twice independently with similar results.

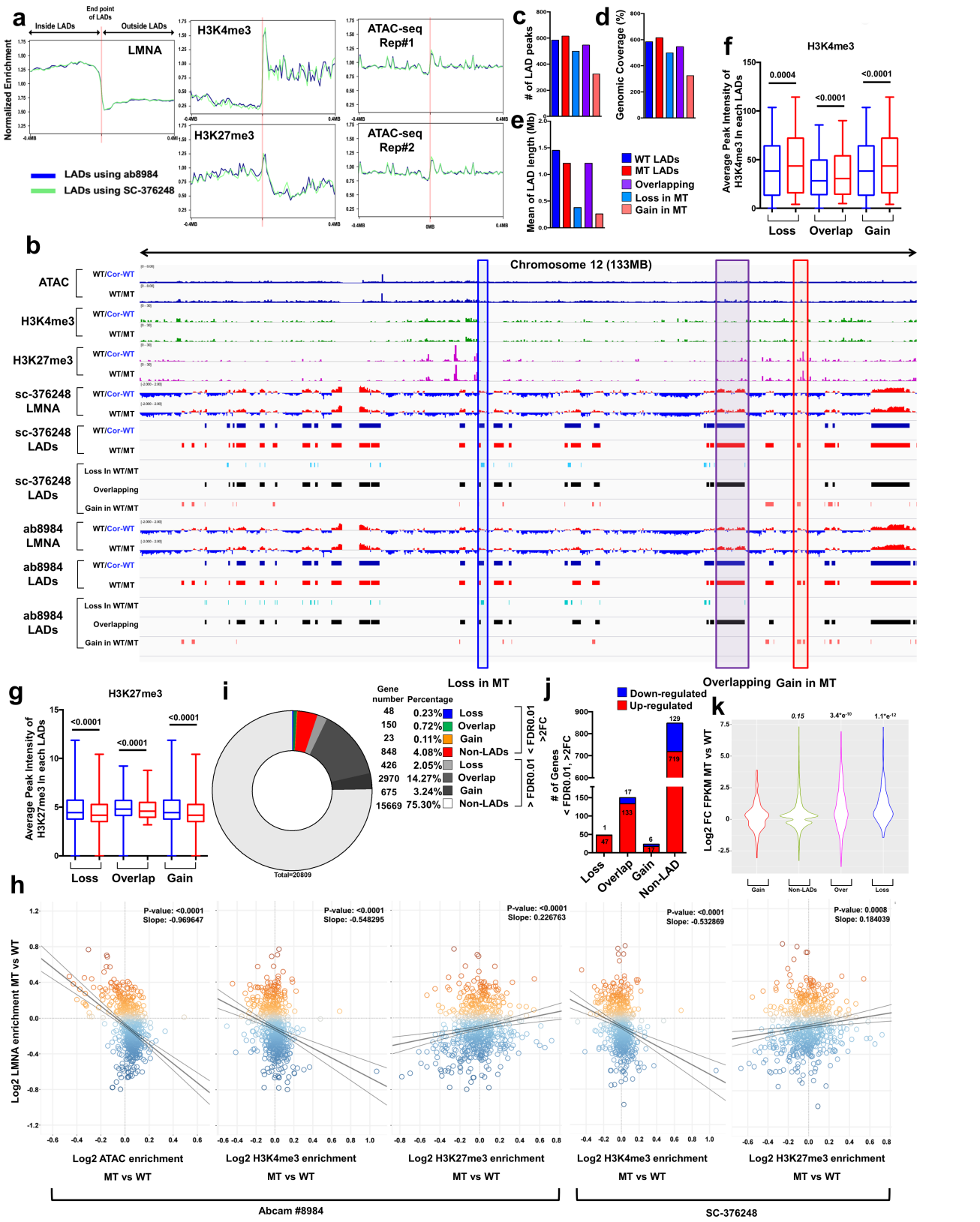


Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Downregulation of mutant mRNA through NMD pathway in *LMNA*-mutant iPSC-CMs. a, Quantification of cells showing abnormal nuclear structures in control and mutant iPSC-CMs. The images were recorded from three differentiation batches. $n = 215$ (WT/cor-WT), $n = 286$ (WT/MUT), $n = 222$ (ins-MUT/MUT) and $n = 280$ (del-KO/MUT). **b**, Representative confocal images of control and mutant lines. Micro-patterned iPSC-CMs were stained with specific antibodies against TNNT2 (red), LMNA (white) and LMNB1 (green). Blue, DAPI. Scale bar, 20 μm. The experiments were repeated three times independently with similar results. **c**, Quantification of cells showing abnormal nuclear structures in control and mutant iPSC-CMs. The images were recorded from three differentiation batches. Data are mean ± s.e.m.; a two-tailed Student's *t*-test was used to calculate *P* values; $n = 3$ (total number of counted cells, 175 (WT/WT) and 203 (WT/ins-MUT)); the value above the line indicates significance. **d**, Immunoblot analysis of lamin A/C levels in control and mutant iPSC-CMs. **e**, Quantification of signal intensity of the lamin A/C band in **d**. Data are mean ± s.e.m.; statistical significance was obtained using one-way ANOVA; values above the line indicate significance; $n = 10$ (WT/WT), $n = 7$ (WT/ins-MUT), $n = 5$ (WT/MUT). **f**, Immunoblot analysis of lamin A/C levels in two different clones of control and mutant iPSC-CMs. Two different antibodies that recognize the N terminus of lamin A/C were used. GAPDH was used as loading control. **g**, Relative mRNA expression of total *LMNA* in control and mutant iPSC-CMs. Data are mean ± s.e.m.; a two-tailed Student's *t*-test was used to calculate *P* values; the value above the line indicates significance; $n = 10$ (WT/WT), $n = 7$ (WT/ins-MUT). **h**, Confirmation of allele-specific primers using plasmid carrying wild-type *LMNA* or mutant *LMNA*. Digital PCR using allele-specific primers detected the ratio of wild-type/mutant *LMNA*, which was consistent with the ratio of wild-type/mutant plasmids. Data are mean ± s.d.; $n = 3$. **i**, Immunoblot analysis of cell lysates from mutant iPSC-CMs treated with emetine and wortmannin. Two different batches of antibodies were used. Red asterisks indicate the truncated lamin A/C with a 14-kDa size. **j**, Immunoblot analysis of cell lysates from mutant iPSC-CMs treated with wortmannin. Three different batches of E-1 antibody detect the N terminus of LMNA and the 131C3 antibody detects the C terminus. **k**, Immunoblot analysis of cell lysates from control iPSC-CMs treated with emetine and wortmannin. The experiments in **f**, **i**–**k** were repeated twice independently with similar results.

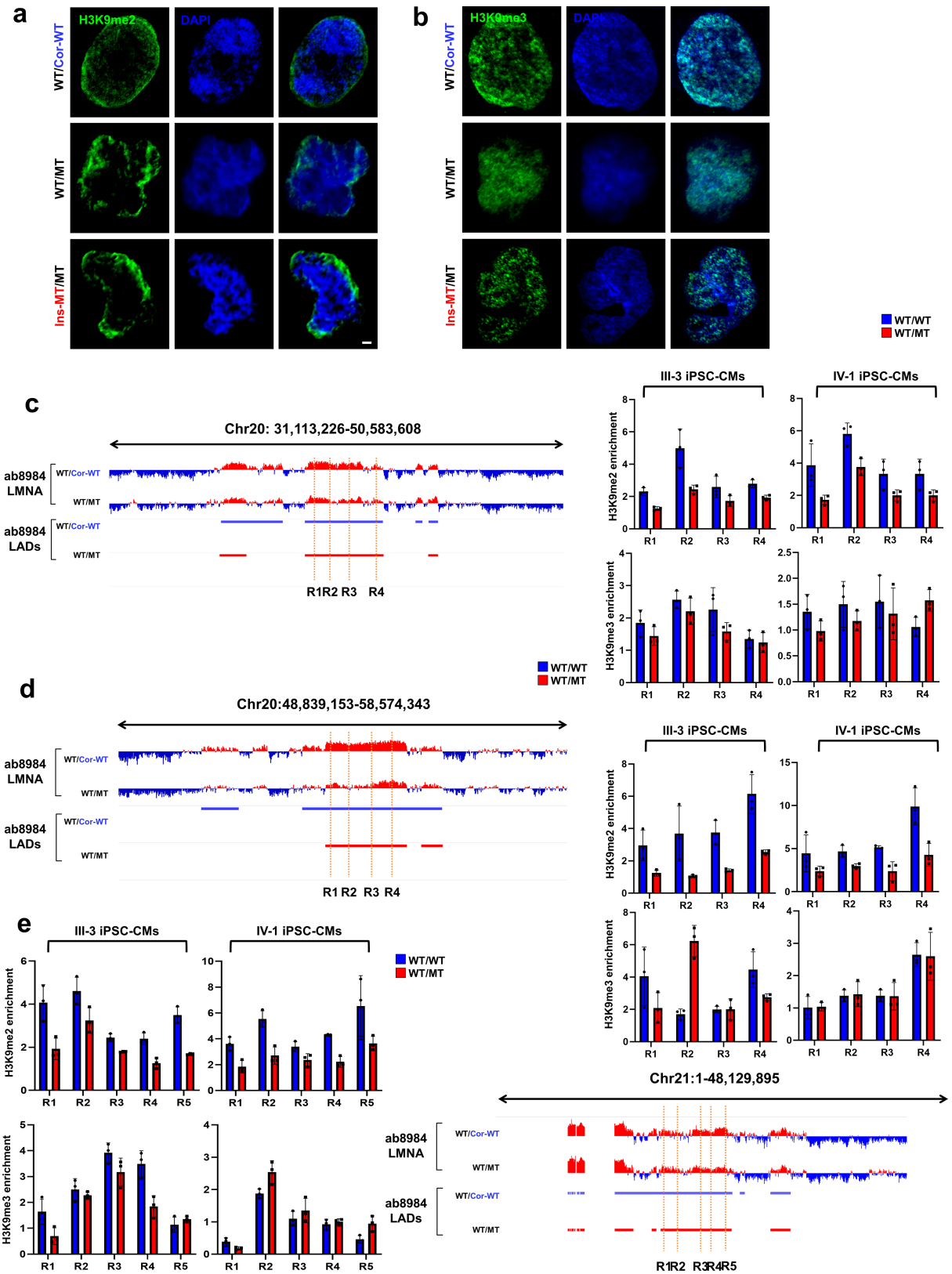


Extended Data Fig. 5 | Lamin A/C Haploinsufficiency results in an abnormal distribution of open chromatin in *LMNA*-mutant iPSC-CMs. **a**–**f**, Representative images and normalized signal intensity of ATAC-see and DAPI of control and mutant iPSC-CMs. Data were obtained from different patient lines, including the uncorrected and isogenic lines of patient III-3 (**a**, **b**); the uncorrected and isogenic line of control IV-1 (**c**, **d**); and the line of patient III-15 (**e**, **f**) for normalized signal intensity of ATAC-see and DAPI. Data are mean ± s.e.m. **g**, Correlation of signal distribution between ATAC-see and DAPI. $n = 42$ (WT/WT), $n = 28$ (WT/cor-WT), $n = 33$ (del-KO/MUT), $n = 32$ (ins-WT/WT), $n = 25$ (WT/MUT) for normalized signal intensity of ATAC-see and DAPI. Data are mean and minimum to maximum; two-tailed Student's *t*-test was used to calculate *P* values. The experiments in **a**, **c**, and **e** were repeated three times independently with similar results.

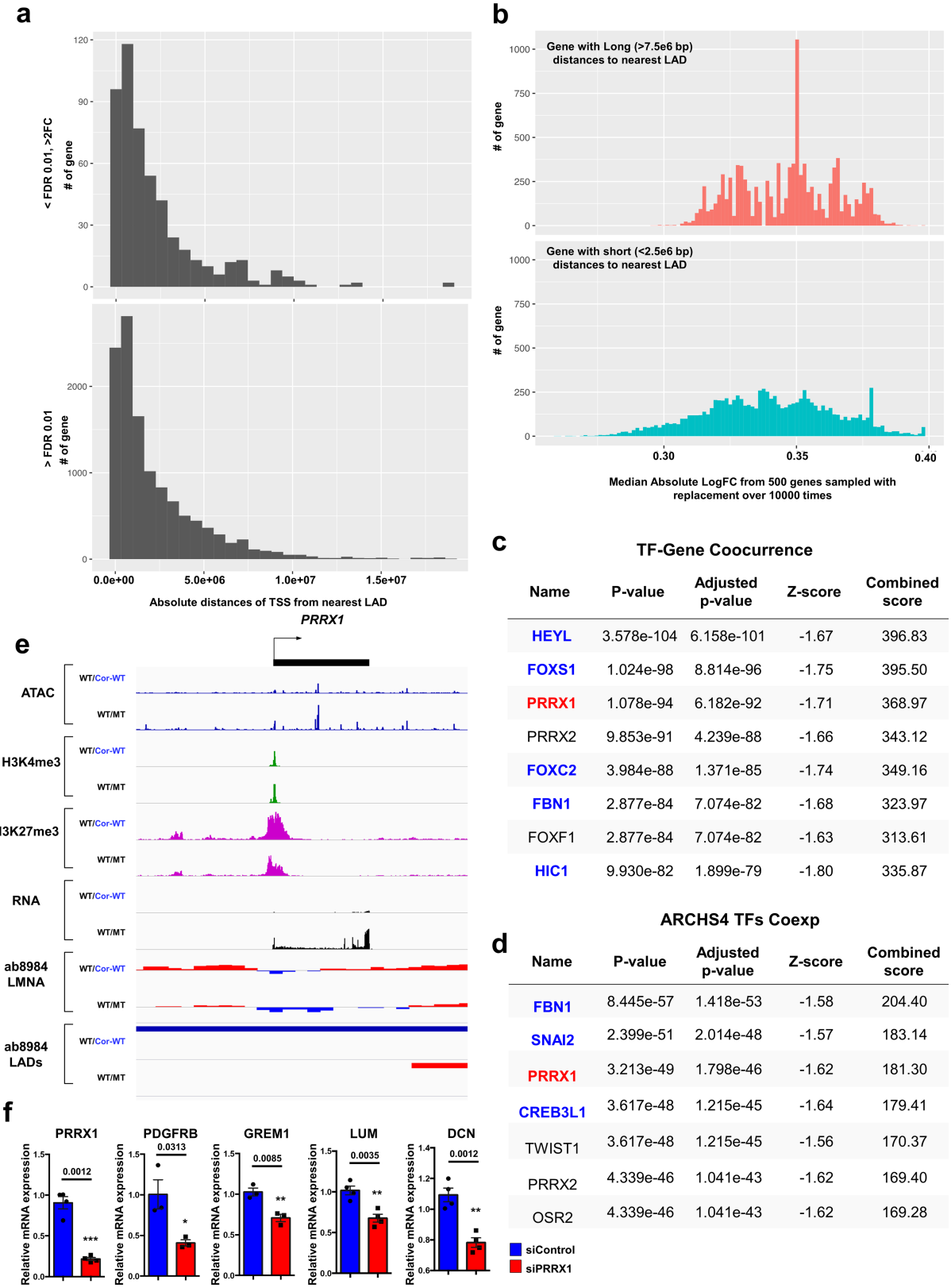


Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Genomic and chromatin features of LADs in control and mutant iPSC-CMs. a, Normalized enrichment of lamin A/C ChIP–seq signals, histone markers (H3K4me3 and H3K27me3) and ATAC-seq signals within ±0.4 Mb of mapped LAD borders. The genomic locations of LADs were obtained from ChIP–seq on lamin A/C using two different antibodies (Abcam 8984, blue line, sc-376248, green line) in control iPSC-CMs (III-3). **b**, Representative images of ChIP–seq, ATAC-seq and RNA-seq of chromosome 12 (133 Mb). The red box shows LADs explicitly called in mutant iPSC-CMs (gain); the purple box shows LADs called in both control and mutant iPSC-CMs (overlapping); the blue box shows LADs explicitly called in control iPSC-CMs (loss). **c**–**e**, Number (**c**), genomic coverage (**d**) and mean of length of LADs (**e**) in control, mutant, gain, overlapping and loss LADs. ChIP–seq on lamin A/C (Abcam 8984) was used for data analysis. **f**, **g**, Average peak intensity of H3K4me3 and H3K27me3 of each LAD. $n = 184$ (loss), $n = 370$ (overlapping), $n = 184$ (gain) for H3K4me3; $n = 273$ (loss), $n = 504$ (overlapping), $n = 273$ (gain) for H3K27me3. Data are mean and minimum to maximum; Wilcoxon matched-pairs signed-rank test was used to calculate *P* values. **h**, Scatter plot of normalized lamin A/C, ATAC and histone marker (H3K4me3 and H3K27me3) enrichment of each LAD. The *y* axis shows the $\log_2$-transformed relative normalized lamin A/C enrichment of each LAD in mutant iPSC-CMs compared to control iPSC-CMs. The *x* axis shows the $\log_2$-transformed relative normalized ATAC and histone marks enrichment of each LAD in mutant iPSC-CMs compared with control iPSC-CMs. Each data point represents one LAD. The statistical significance was obtained using one-way ANOVA; $n = 587$ for sc-376248 and $n = 585$ for Abcam 8984. **i**, Percentage of differentially expressed genes in mutant iPSC-CMs compared to control iPSC-CMs. **j**, Number of differentially expressed genes located in mutant iPSC-CMs compared to control iPSC-CMs. (FDR-adjusted $P < 0.01$; $\log_2$-transformed fold change in expression of >1 or <-1). **k**, Distribution of $\log_2$-transformed fold change in FPKM in control and mutant iPSC-CMs. A non-parametric Kruskal–Wallis (testing for two-sided differences) followed by Dunn's post hoc test was used to adjust for multiple comparisons; $n = 266$ (gain), $n = 8171$ (non-LADs), $n = 835$ (overlapping), $n = 206$ (loss).

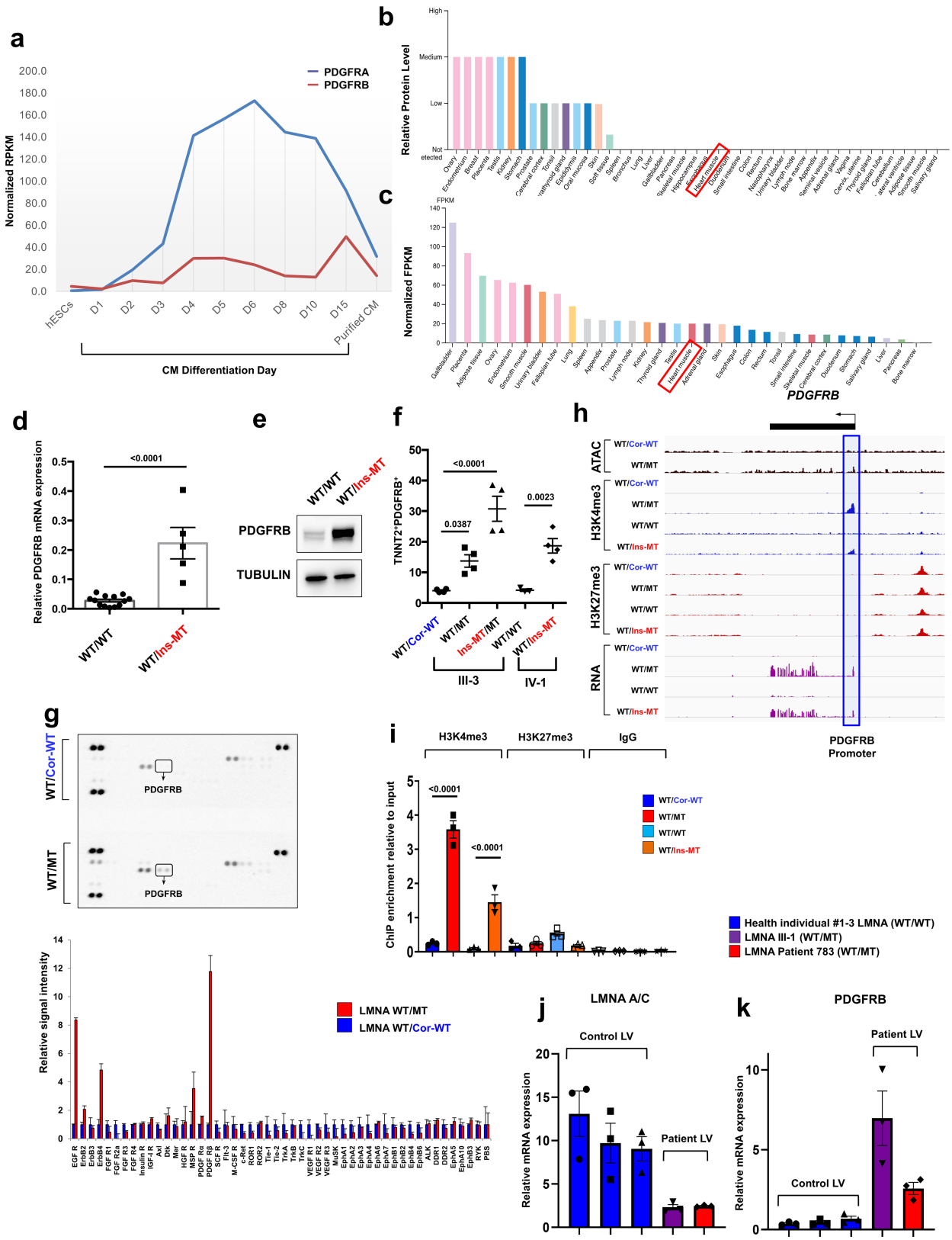


Extended Data Fig. 7 | Abnormal distribution of H3K9 methylation in mutant iPSC-CMs. a, **b**, Representative images of immunofluorescence staining of control mutant iPSC-CMs. iPSC-CMs were stained with specific antibodies against H3K9me2 or H3K9me3 (green). Blue, DAPI. Scale bar, 1,000 nm. The experiments were repeated three times independently with similar results. **c**–**e**, Representative images of lamin A/C enrichment and LAD distribution of ChIP–seq data. ChIP–qPCR analysis of H3K9me2 and H3K9me3 enrichment on LAD regions. Data are mean ± s.d.; $n = 3$.



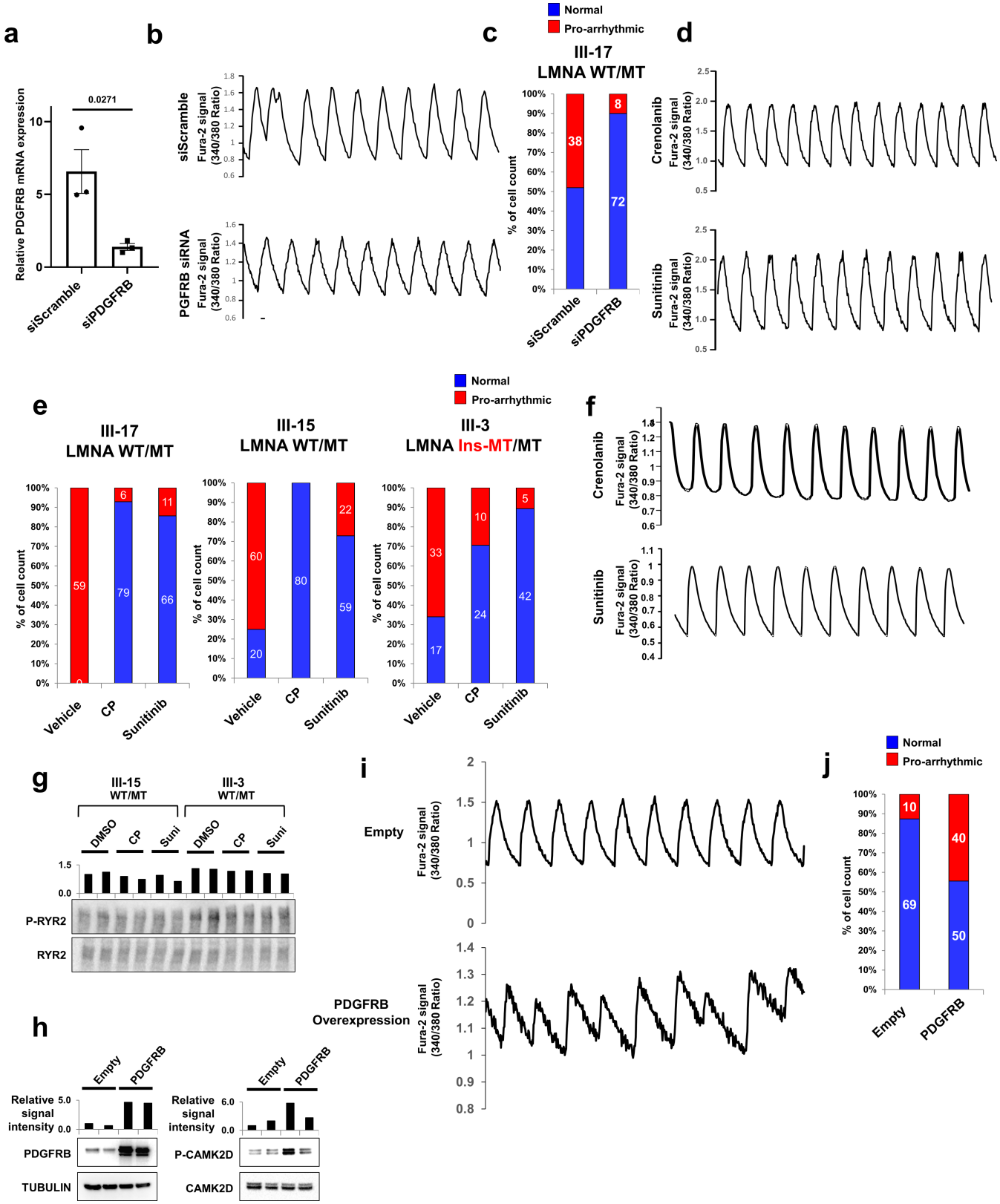
Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Transcription factors altered by lamin A/C haploinsufficiency contribute to the activation of genes located outside LADs. a, Distribution of absolute distances to the nearest LAD (by nucleotide distance) from the transcription start site of genes that are differentially expressed (top) or that show no significant difference in expression between mutant and control iPSC-CMs (bottom). **b**, Distribution of median absolute $\log_2$-transformed change in expression of genes with relatively long ($>7.5 \times 10^6$ bp) distances to the nearest LAD (top) and genes with relatively short ($<2.5 \times 10^6$ bp) distances to the nearest LAD (bottom). In each category, 500 genes were sampled with replacement over 10,000 times. **c**, **d**, Co-occurrence analyses of transcription factors and genes and coexpression analyses of ARCHS4 transcription factors of differentially expressed genes located in non-LADs. Genes located in non-LADs are shown in blue; genes with no significant difference in gene expression between control and mutant iPSC-CMs are shown in black; genes located in LADs and highly expressed in mutant iPSC-CMs compared with control iPSC-CMs are shown in red. Top 200 differentially expressed genes located in non-LADs were used for the analysis. **e**, Representative images of ChIP–seq, ATAC-seq and RNA-seq of the genomic region of *PRRX1*. **f**, Relative mRNA expression of *PRRX1*, *PDGFRB*, *GREM1*, *LUM* and *DCN* in mutant iPSC-CMs treated with scramble or *PRRX1* siRNA. Data are mean ± s.e.m.; a two-tailed Student's *t*-test was used to calculate *P* values; $n = 3$ (*PDGFRB* and *GREM1*), $n = 4$ (*DCN*, *LUM* and *PRRX1*); values above the lines show significance.



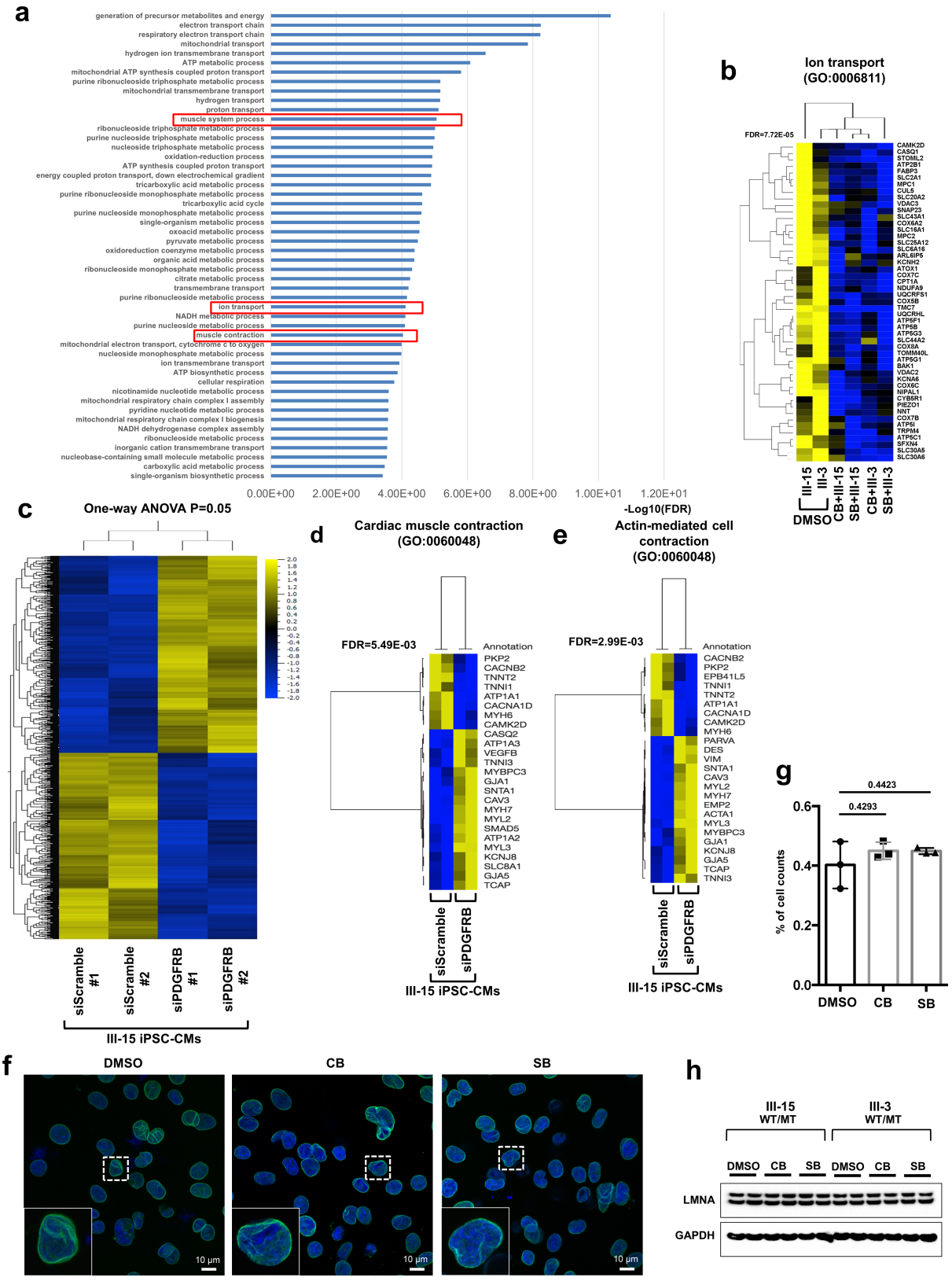
Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | PDGFRB is upregulated in *LMNA*-mutant iPSC-CMs. a, Expression levels of *PDGFRA* and *PDGFRB* during the human iPSC-CM differentiation process. The data were adapted from previously published data (GSE76523). **b**, **c**, Protein and RNA levels of *PDGFRB* in human tissues. The data were adapted from the Human Protein Atlas Database[32] v.18.1 (data available from http://www.proteinatlas.org/). **d**, qPCR analysis of *PDGFRB* expression in *LMNA*-mutant and control iPSC-CMs. Data are mean ± s.e.m.; a two-tailed Student's *t*-test was used to calculate *P* values; $n = 13$ (WT/WT), $n = 5$ (WT/ins-MUT); the value above the line shows significance. **e**, Immunoblot analysis of PDGFRB protein levels in control versus mutant iPSC-CMs. GAPDH was used as loading control. The experiments were repeated twice independently with similar results. **f**, Flow cytometry analysis of $TNNT2^{+}PDGFRB^{+}$ cells in control and mutant iPSC-CMs. $n = 4$. **g**, Kinase array of control and mutant iPSC-CMs. Fifty different protein kinases were presented in each chip. Top, raw images of the blotting membrane. Two dots carried the same antibody in technical duplicates. Bottom, quantification of the signal intensity of each spot. **h**, Representative images of ChIP–seq, ATAC-seq and RNA-seq on the genomic regions of *PDGFRB*. The promoter region of *PDGFRB* is highlighted by a blue box. **i**, ChIP–qPCR of H3K4me3 and H3K27me3 enrichment at the promoter region of *PDGFRB* in control and mutant iPSC-CMs. $n = 3$. **j**, **k**, qPCR analysis of *LMNA* and *PDGFRB* expression levels in left ventricular heart tissue from health controls ($n = 3$) and patients with LMNA-related DCM ($n = 2$). Data are mean ± s.e.m. The kinase data in **g** were repeated twice independently with similar results. **f**, **i**, Data are mean ± s.e.m.; statistical significance was obtained using one-way ANOVA; values above the lines show significance.



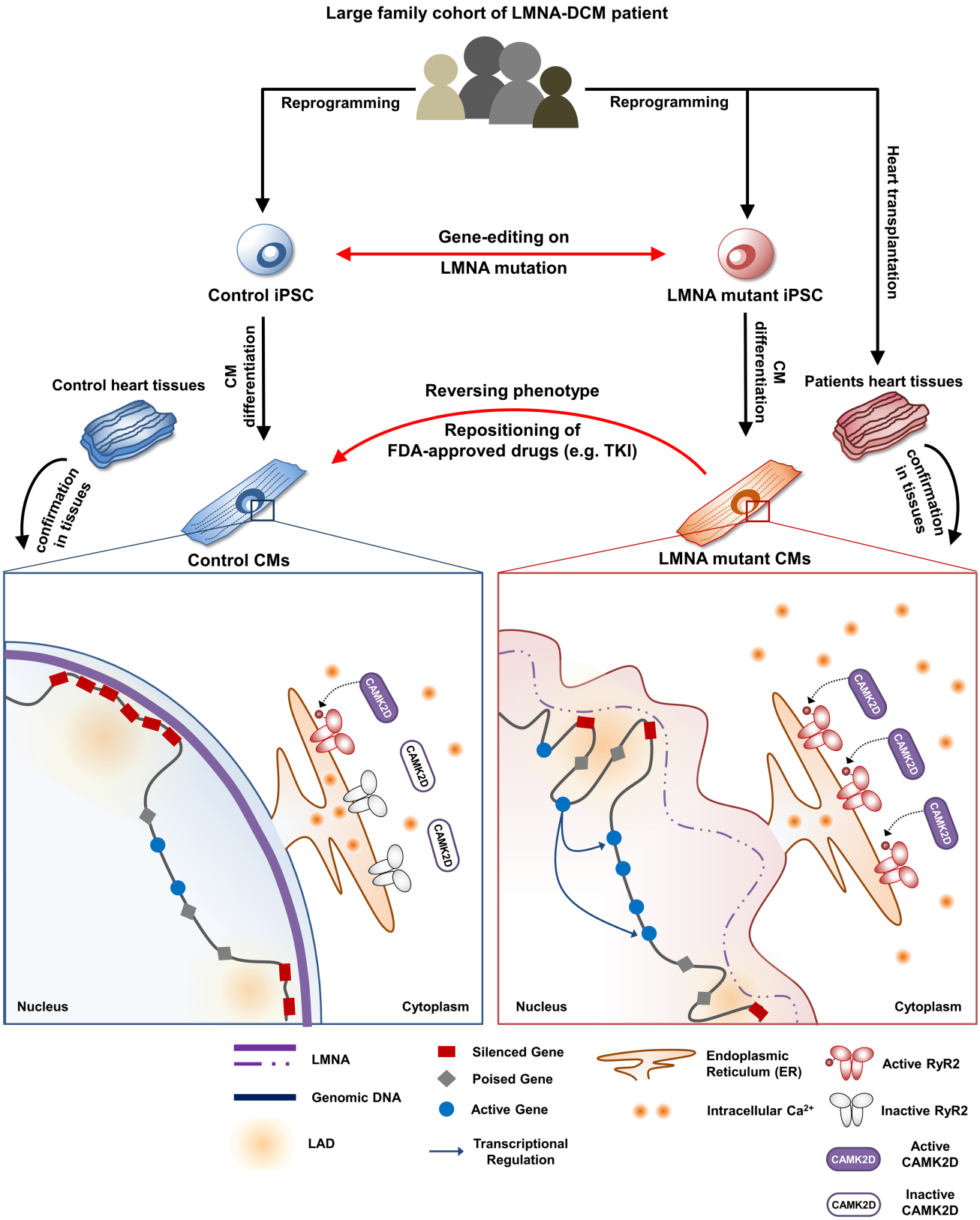
Extended Data Fig. 10 | See next page for caption.

Extended Data Fig. 10 | Arrhythmic phenotype in mutant iPSC-CMs is dependent on the activation of the PDGFRB pathway. a, qPCR analysis of *PDGFRB* expression levels in mutant iPSC-CMs (WT/MUT) treated with scramble or *PDGFRB* siRNAs. The cells were treated with siRNAs for 48 h. Data are mean ± s.e.m.; a two-tailed Student's *t*-test was used to calculate *P* values; $n = 3$; the value above the line indicates significance. **b**, Representative Ca^{2+} transients of mutant iPSC-CMs (III-17 WT/MUT) treated with scramble siRNA or *PDGFRB* siRNA. **c**, Quantification of the number of cells that exhibited arrhythmic waveforms in **b**. **d**, Representative Ca^{2+} transients of mutant iPSC-CMs treated with PDGRB inhibitors, crenolanib (100 nM) and sunitinib (500 nM), for 24 h. All traces were recorded for 20 s. **e**, Quantification of mutant iPSC-CMs (III-17, III-15 and III-3) that exhibited arrhythmic waveforms with or without the treatment of PDGRB inhibitors, crenolanib (100 nM) and sunitinib (500 nM), for 24 h. **f**, Representative Ca^{2+} transients of mutant iPSC-CMs (III-17 WT/MUT) treated with PDGFRB inhibitors. **g**, Immunoblot analysis of pRYR2 and RYR2 protein levels with treatment of DMSO, crenolanib or sunitinib. The data were repeated twice independently with similar results. **h**, Immunoblot analysis of PDGFRB, tubulin, pCAMK2D and CAMK2D protein levels in control iPSC-CMs expressing empty and *PDGFRB* constructs. The signal intensity of the PDGFRB (left) and p-CAMK2D (right) is shown. The experiments were repeated twice independently with similar results. **i**, Representative Ca^{2+} transients of iPSC-CMs expressing empty and *PDGFRB* constructs. **j**, Quantification of arrhythmic waveforms of iPSC-CMs in **i**. The Ca^{2+} transients in **b**, **d**, **f** and **i** were repeated as described in **c**, **e** and **j** independently with similar results.



Extended Data Fig. 11 | See next page for caption.

Extended Data Fig. 11 | Gene-expression profile of PDGFRB inhibition in *LMNA*-mutant iPSC-CMs. a, GO analysis of downregulated genes ($n = 352$) in *LMNA*-mutant iPSC-CMs treated with PDGFRB inhibitors, crenolanib (100 nM) and sunitinib (500 nM), for 24 h. **b**, Heat map of the expression profile of the gene set related to the GO function of ion transport. The FDR-adjusted *P* values were obtained using the GO enrichment analysis tool. **c**, Hierarchical clustering of AmpliSeq RNA-seq data using one-way ANOVA ($P = 0.05$; $n = 230$). Two different siRNAs against *PDGFRB* and a scramble siRNA were used in *LMNA*-mutant iPSC-CMs (III-15 WT/MUT). **d**, **e**, Heat map of expression profile of gene ($n = 25$) sets related with the GO function of cardiac muscle contraction (**d**) and actin-mediated cell contraction (**e**). The FDR-adjusted *P* values were obtained using the GO enrichment analysis tool. **f**, No significant changes in abnormal nuclear structures of mutant iPSC-CMs by inhibition of PDGFRB were found. Representative images of mutant iPSC-CMs treated with PDGFRB inhibitors, crenolanib (100 nM) and sunitinib (500 nM), for 24 h. iPSC-CMs were stained with specific antibodies against LMNB1 (green). Blue, DAPI. Scale bars, 10 μm. The experiments were repeated three times independently with similar results. **g**, Quantification of cells showing abnormal nuclear structures in mutant iPSC-CMs treated with PDGFRB inhibitors. The images were recorded from three differentiation batches. $n = 90$ (DMSO), $n = 69$ (crenolanib), $n = 79$ (sunitinib). Data are mean ± s.e.m.; statistical significance was analysed using one-way ANOVA; values above the lines indicate significance. **h**, Immunoblot analysis of lamin A/C and GAPDH protein levels in mutant iPSC-CMs treated with PDGFRB inhibitors. CB, crenolanib; SB, sunitinib. The experiments were repeated twice independently with similar results.



Extended Data Fig. 12 | See next page for caption.

Extended Data Fig. 12 | Proposed disease model of LMNA-related DCM. We recruited a large family cohort with DCM and generated patient-specific iPSCs from several patients ($n = 5$) and healthy individuals ($n = 2$). We next used gene-edited isogenic iPSC lines ($n = 4$) and patient heart tissues to address the question why patients with LMNA-related DCM have increased manifestation of cardiac arrhythmias. The electrophysiological studies of mutant iPSC-CMs demonstrated that a mutation in *LMNA* was the cause of the increased arrhythmogenicity in LMNA-mutant iPSC-CMs. We also found that the LMNA mutation caused lamin A/C haploinsufficiency, which led to abnormal calcium homeostasis in mutant iPSC-CMs through upregulation of calcium-handling genes. Whole-transcriptome profiling (RNA-seq) further demonstrated an abnormal activation of the PDGF pathway in mutant iPSC-CMs. The inhibition of the PDGF signalling pathway by treatment with siRNA or FDA-approved inhibitors, such as sunitinib and crenolanib, could reverse the arrhythmic phenotype of *LMNA*-mutant iPSC-CMs. Cross-analysis of ChIP–seq, ATAC-seq and RNA-seq data revealed a possible underlying mechanism that lamin A/C haploinsufficiency could disrupt global chromatin conformation, resulting in abnormal gene expression in mutant iPSC-CMs. These findings were further corroborated by studies in cardiac tissues from healthy individuals and patients with LMNA-related DCM, thus validating a novel mechanism of LMNA-related DCM pathogenesis both in vitro and in vivo.

nature research

Corresponding author(s): Joseph C. Wu

Last updated by author(s): May 2, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	-qPCR analysis: CFX Maestro™ Software (Bio-rad) -Imaging process for ATAC-See: ZEN imaging software (Zeiss) -Imaging process for kinase array: Image Studio Lite (LI-COR) -Imaging process for immunoblot: Image Lab (Bio-rad) -AmpliSeq Transcriptome analysis: Omics Explorer version 3.2 software (Qlucore) -Ca2+ imaging: NIS Elements AR software
Data analysis	- [RNA-seq] Base quality of raw reads: FastQC 0.11.4 - [RNA-seq] aligned the reads to the human reference genome (hg19): STAR 2.5.1b - [RNA-seq] calculating fragments per kilobase per million aligned reads (FPKM): Cufflinks 2.2.1. - [ATAC-seq] ATAC-seq reads were mapped to hg19 genome: bowtie 2 - [ATAC-seq] Following QC to remove duplicate reads, average read intensities were calculated: deepTools and R/Bioconductor (v.3.2.1) - [ChIP-seq] AQUAS pipeline from the Kundaje lab at Stanford University (https://github.com/kundajelab/chip-seq-pipeline2) - [ChIP-seq] duplicate reads were removed: MarkDuplicates from Picard Tools (v2.17.3) - [ChIP-seq] LMNA data were analyzed: Enriched Domain Detector (v1.0) with an 11 Kb bin size, gap penalty of 5, and FDR significance threshold of 0.05. - [ChIP-seq] LAD gain, loss, and intersection were found: bedtools (v2.27.1)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

GSE118885

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to predetermine sample size. We included all patients who provided consent in our study.
Data exclusions	No data excluded from the analysis.
Replication	For each experiment, all attempts at replication were successful.
Randomization	The experiments were not randomized. We allocated our samples into two groups based on genotype of LMNA gene.
Blinding	The investigators who performed electro-physiological test, Ca2+ imaging analysis and measuring abnormal nuclear structure were blinded to group allocation during experiments and data collection.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☐	☒ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

[SSEA4]
-Company (catalog number): R&D systems (MAB1435)
-Application: Immunostaining
-Source: Monoclonal Mouse IgG3 Clone # MC-813-70
[Oct-3/4]
-Company (catalog number): Santa Cruz Biotechnology (sc-5279)
-Application: Immunostaining
-Source: mouse monoclonal IgG2b # C-10
[NANOG]
-Company (catalog number): Santa Cruz Biotechnology (sc-33760)
-Application: Immunostaining
-Source: rabbit polyclonal IgG # M-149
[SOX2]
-Company (catalog number): R&D systems (MAB2018)

-Application: Immunostaining
-Source: Monoclonal Mouse IgG2A Clone # 245610
[Cardiac troponin T]
-Company (catalog number): Abcam (ab45932)
-Application: Immunostaining
-Source: Rabbit polyclonal IgG
[Cardiac troponin T]
-Company (catalog number): Abcam (ab8295)
-Application: Immunostaining
-Source: Mouse monoclonal [1C11]
[α-Actinin (Sarcomeric)]
-Company (catalog number): Sigma-Aldrich (A7811)
-Application: Immunostaining
-Source: Mouse monoclonal EA-53
[LMNA]
-Company (catalog number): abcam (ab8980)
-Application: Immunoblotting, ChIP-seq
-Source: Mouse monoclonal [133A2]
[LMNA]
-Company (catalog number): Santa Cruz Biotechnology (sc-376248)
-Application: Immunoblotting, Immunostaining, ChIP-seq
-Source: mouse monoclonal IgG1 (E-1)
[LMNA]
-Company (catalog number): abcam (ab8984)
-Application: Immunoblotting
-Source: Mouse monoclonal [131C3]
[LMNB1]
-Company (catalog number): abcam (ab16048)
-Application: Immunostaining
-Source: Rabbit polyclonal
[CaMKII delta]
-Company (catalog number): Abcam (ab181052)
-Application: Immunoblotting
-Source: Mouse monoclonal [EPR13095]
[Phospho-CaMKII delta]
-Company (catalog number): Abcam (ab32678)
-Application: Immunoblotting
-Source: Rabbit polyclonal
[PDGF Receptor beta]
-Company (catalog number): Abcam (ab32570)
-Application: Immunoblotting
-Source: Rabbit monoclonal [Y92]
[Ryanodine Receptor]
-Company (catalog number): Abcam (ab2868)
-Application: Immunoblotting
-Source: Mouse monoclonal [34C]
[SSEA4]
-Ref in manufacturer's web site: Shevinsky, L.H. et al. (1982) Cell 30:697. Kannagi, R. et al. (1983) EMBO J. 2:2355.
[Oct-3/4]
-Ref in manufacturer's web site: PMID: # 27797132, PMID: # 28159471, PMID: # 27876565
[NANOG]
-Ref in manufacturer's web site: PMID: # 24550733, PMID: # 24619130, PMID: # 24380431
[SOX2]
-Ref in manufacturer's web site: Graham, V. et al. (2003) Neuron 39:749., Avilion, A.A. et al. (2003) Genes Dev. 17:126.
[Cardiac troponin T]
-Abcam (ab45932)- Ref in manufacturer's web site: PubMed: 27672365, PubMed: 27226619
-Abcam (ab8295)- -Ref in manufacturer's web site: PubMed: 28487655, PubMed: 28490375.
[α-Actinin (Sarcomeric)]
-Ref in manufacturer's web site: PMID 19668186, PMID 22020047
[LMNA]
-Abcam (ab8980)- Ref in manufacturer's web site: PubMed: 29545600, PubMed: 28737169
-SCBT (sc-376248)- Ref in manufacturer's web site: PMID: # 29684352, PMID: # 29436586
- Abcam (ab8984)- Ref in manufacturer's web site: PubMed: 29580221, PubMed: 29659505
[CaMKII delta]
--Ref in manufacturer's web site: PubMed: 27084844
[Phospho-CaMKII delta]
--Ref in manufacturer's web site: PubMed: 29482582, PubMed: 29593308
[PDGF Receptor beta]
-Ref in manufacturer's web site: PubMed: 28423550, PubMed: 28230073
[Ryanodine Receptor]
-Ref in manufacturer's web site: PubMed: 26301072, PubMed: 25775120
[LMNB1]
-Ref in manufacturer's web site: PubMed: 29308302, PubMed: 29335436
[H3K9me2]
-Company (catalog number): Abcam (ab1220), Active Motif (#39239)

-Application: Immunostaining, ChIP-qPCR
-Ref: PMID 29033129
[H3K9me3]
-Company (catalog number): Active Motif (#61013)
-Application: Immunostaining, ChIP-qPCR
-Ref: PMID:30143619

Validation

[SSEA4]
-Company (catalog number): R&D systems (MAB1435)
-Application: Immunostaining
-Source: Monoclonal Mouse IgG3 Clone # MC-813-70
[Oct-3/4]
-Company (catalog number): Santa Cruz Biotechnology (sc-5279)
-Application: Immunostaining
-Source: mouse monoclonal IgG2b # C-10
[NANOG]
-Company (catalog number): Santa Cruz Biotechnology (sc-33760)
-Application: Immunostaining
-Source: rabbit polyclonal IgG # M-149
[SOX2]
-Company (catalog number): R&D systems (MAB2018)
-Application: Immunostaining
-Source: Monoclonal Mouse IgG2A Clone # 245610
[Cardiac troponin T]
-Company (catalog number): Abcam (ab45932)
-Application: Immunostaining
-Source: Rabbit polyclonal IgG
[Cardiac troponin T]
-Company (catalog number): Abcam (ab8295)
-Application: Immunostaining
-Source: Mouse monoclonal [1C11]
[α-Actinin (Sarcomeric)]
-Company (catalog number): Sigma-Aldrich (A7811)
-Application: Immunostaining
-Source: Mouse monoclonal EA-53
[LMNA]
-Company (catalog number): abcam (ab8980)
-Application: Immunoblotting, ChIP-seq
-Source: Mouse monoclonal [133A2]
[LMNA]
-Company (catalog number): Santa Cruz Biotechnology (sc-376248)
-Application: Immunoblotting, Immunostaining, ChIP-seq
-Source: mouse monoclonal IgG1 (E-1)
[LMNA]
-Company (catalog number): abcam (ab8984)
-Application: Immunoblotting
-Source: Mouse monoclonal [131C3]
[LMNB1]
-Company (catalog number): abcam (ab16048)
-Application: Immunostaining
-Source: Rabbit polyclonal
[CaMKII delta]
-Company (catalog number): Abcam (ab181052)
-Application: Immunoblotting
-Source: Mouse monoclonal [EPR13095]
[Phospho-CaMKII delta]
-Company (catalog number): Abcam (ab32678)
-Application: Immunoblotting
-Source: Rabbit polyclonal
[PDGF Receptor beta]
-Company (catalog number): Abcam (ab32570)
-Application: Immunoblotting
-Source: Rabbit monoclonal [Y92]
[Ryanodine Receptor]
-Company (catalog number): Abcam (ab2868)
-Application: Immunoblotting
-Source: Mouse monoclonal [34C]
[SSEA4]
-Ref in manufacturer's web site: Shevinsky, L.H. et al. (1982) Cell 30:697. Kannagi, R. et al. (1983) EMBO J. 2:2355.
[Oct-3/4]
-Ref in manufacturer's web site: PMID: # 27797132, PMID: # 28159471, PMID: # 27876565
[NANOG]
-Ref in manufacturer's web site: PMID: # 24550733, PMID: # 24619130, PMID: # 24380431
[SOX2]
-Ref in manufacturer's web site: Graham, V. et al. (2003) Neuron 39:749., Avilion, A.A. et al. (2003) Genes Dev. 17:126.

[Cardiac troponin T]
-Abcam (ab45932)- Ref in manufacturer's web site: PubMed: 27672365, PubMed: 27226619
-Abcam (ab8295)- -Ref in manufacturer's web site: PubMed: 28487655, PubMed: 28490375.
[α-Actinin (Sarcomeric)]
-Ref in manufacturer's web site: PMID 19668186, PMID 22020047
[LMNA]
-Abcam (ab8980)- Ref in manufacturer's web site: PubMed: 29545600, PubMed: 28737169
-SCBT (sc-376248)- Ref in manufacturer's web site: PMID: # 29684352, PMID: # 29436586
- Abcam (ab8984)- Ref in manufacturer's web site: PubMed: 29580221, PubMed: 29659505
[CaMKII delta]
--Ref in manufacturer's web site: PubMed: 27084844
[Phospho-CaMKII delta]
--Ref in manufacturer's web site: PubMed: 29482582, PubMed: 29593308
[PDGF Receptor beta]
-Ref in manufacturer's web site: PubMed: 28423550, PubMed: 28230073
[Ryanodine Receptor]
-Ref in manufacturer's web site: PubMed: 26301072, PubMed: 25775120
[LMNB1]
-Ref in manufacturer's web site: PubMed: 29308302, PubMed: 29335436
[H3K9me2]
-Company (catalog number): Abcam (ab1220), Active Motif (#39239)
-Application: Immunostaining, ChIP-qPCR
-Ref: PMID 29033129
[H3K9me3]
-Company (catalog number): Active Motif (#61013)
-Application: Immunostaining, ChIP-qPCR
-Ref: PMID:30143619

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	-We obtained dermal fibroblasts or PBMCs from the patients and generated patient specific iPSC lines using Co-MIPs (doi:10.1038/srep08081) or sendai virus method (ThermoFisher, CytoTune™-iPS 2.0 Sendai Reprogramming Kit; A16517). -H7 hESCs line were obtained from WiCell (WAe007-A). -The iPSC-derived cardiomycytes were generated by previous protocol (doi:10.3791/52628).
Authentication	-Immunofluresence assay of each iPSC line was performed to check the expression of stem cell markers such as NANOG, POU5F1 and SOX2. -SNP karyotyping was tested through HuCytoSNP-12 chip (Illumina), and CNV and SNP visualization was performed using KaryoStudio v1.4 (Illumina).
Mycoplasma contamination	We confirmed that all cell lines were negative for mycoplasma contamination using MycoAlert™ PLUS Mycoplasma Detection Kit (Lonza, LT07-705).
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

Human research participants

Policy information about studies involving human research participants

Population characteristics	The detail information of patients was described in Extended Data Figure 1b. Patients III-1, III-3, and III-9 (age 57, 60, and 67, respectively) carried the c.349_350insG frame shift mutation on LMNA gene initially presented with atrial fibrillation that progressed to atrioventricular block and ventricular tachycardia. These patients eventually required implantable cardioverter defibrillators (ICDs) and later developed DCM. The other two carriers (III-15 and III-17; ages 38 and 45, respectively) were younger and had exhibited paroxysmal atrial fibrillation prior to the beginning of the study. III-1, III-3, III-9, IV-1, and IV-2 individuals are male. III-15 and III-17 individuals are female.
Recruitment	The fibroblasts, PBMCs, and heart tissues were obtained from patients using IRB-approved protocol at Stanford University (Protocol ID 17576 and 29904). Informed consents were obtained from all patients who were included in our study.
Ethics oversight	IRB-approved protocols at Stanford University (Protocol ID 17576 and 29904)

Note that full information on the approval of the study protocol must also be provided in the manuscript.

ChIP-seq

Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links
May remain private before publication.

Gene Expression Omnibus (GEO): GSE118885

Files in database submission

<Processed>
RNA_PT3WT_Rep1Aligned.out.sorted.q255.bw
RNA_PT3WT_Rep2Aligned.out.sorted.q255.bw
RNA_PT3MT_Rep1Aligned.out.sorted.q255.bw
RNA_PT3MT_Rep2Aligned.out.sorted.q255.bw
RNA_PT5WT_Rep1Aligned.out.sorted.q255.bw
RNA_PT5WT_Rep2Aligned.out.sorted.q255.bw
RNA_PT5MT_Rep1Aligned.out.sorted.q255.bw
RNA_PT5MT_Rep2Aligned.out.sorted.q255.bw
PT3WT_ATAC_USPD16084651_HJ35WBBXX_L8_1.trim.PE2SE.nodup.tn5_pooled.pf.fc.signal.bigwig
PT3MT_ATAC_USPD16084652_HJ35WBBXX_L8_1.trim.PE2SE.nodup.tn5_pooled.pf.fc.signal.bigwig
H3K4_PT3WT.paired.PE2SE.nodup.tagAlign_x_Input_1_R1.paired.PE2SE.nodup.tagAlign.fc.signal.bw
H3K4_PT3MT.paired.PE2SE.nodup.tagAlign_x_Input_2_R1.paired.PE2SE.nodup.tagAlign.fc.signal.bw
H3K27_PT3MT.paired.PE2SE.nodup.tagAlign_x_Input_2_R1.paired.PE2SE.nodup.tagAlign.fc.signal.bw
H3K27_PT3WT.paired.PE2SE.nodup.tagAlign_x_Input_1_R1.paired.PE2SE.nodup.tagAlign.fc.signal.bw
PT3WT_EDD_abLMNA_vs_Input_peaks.bed
PT3MT_EDD_abLMNA_vs_Input_peaks.bed
PT3WT_EDD_SCLMNA_vs_Input_peaks.bed
PT3MT_EDD_SCLMNA_vs_Input_peaks.bed

<Raw Data>
RNA_3_1_1.fq.gz
RNA_3_1_2.fq.gz
RNA_3_2_1.fq.gz
RNA_3_2_2.fq.gz
RNA_3W_1_1.fq.gz
RNA_3W_1_2.fq.gz
RNA_3W_2_1.fq.gz
RNA_3W_2_2.fq.gz
RNA_5_1_1.fq.gz
RNA_5_1_2.fq.gz
RNA_5_2_1.fq.gz
RNA_5_2_2.fq.gz
RNA_5MT_1_1.fq.gz
RNA_5MT_1_2.fq.gz
RNA_5MT_2_1.fq.gz
RNA_5MT_2_2.fq.gz
ATAC_A1_USPD16084651_HJ35WBBXX_L8_1.fq.gz
ATAC_A1_USPD16084651_HJ35WBBXX_L8_2.fq.gz
ATAC_B1_USPD16084655_HJ35WBBXX_L8_1.fq.gz
ATAC_B1_USPD16084655_HJ35WBBXX_L8_2.fq.gz
ATAC_A2_USPD16084652_HJ35WBBXX_L8_1.fq.gz
ATAC_A2_USPD16084652_HJ35WBBXX_L8_2.fq.gz
ATAC_B2_USPD16084656_HJ35WBBXX_L8_1.fq.gz
ATAC_B2_USPD16084656_HJ35WBBXX_L8_2.fq.gz
Input_1_USPD16084627_HJ3FYBBXX_L1_1.fq.gz
Input_1_USPD16084627_HJ3FYBBXX_L1_2.fq.gz
Input_2_USPD16084628_HJ3FYBBXX_L1_1.fq.gz
Input_2_USPD16084628_HJ3FYBBXX_L1_2.fq.gz
H3K4_13_USPD16084639_HJ3FYBBXX_L3_1.fq.gz
H3K4_13_USPD16084639_HJ3FYBBXX_L3_2.fq.gz
H3K4_14_USPD16084640_HJ3FYBBXX_L3_1.fq.gz
H3K4_14_USPD16084640_HJ3FYBBXX_L3_2.fq.gz
H3K27_17_USPD16084643_HJ3FYBBXX_L3_1.fq.gz
H3K27_17_USPD16084643_HJ3FYBBXX_L3_2.fq.gz
H3K27_18_USPD16084644_HJ3FYBBXX_L3_1.fq.gz
H3K27_18_USPD16084644_HJ3FYBBXX_L3_2.fq.gz
ab_9_USPD16084635_HJ3FYBBXX_L2_1.fq.gz
ab_9_USPD16084635_HJ3FYBBXX_L2_2.fq.gz
ab_10_USPD16084636_HJ3FYBBXX_L2_1.fq.gz
ab_10_USPD16084636_HJ3FYBBXX_L2_2.fq.gz
SC_5_USPD16084631_HJ3FYBBXX_L2_1.fq.gz

SC_5_USPD16084631_HJ3FYBBXX_L2_2.fq.gz
SC_6_USPD16084632_HJ3FYBBXX_L2_1.fq.gz
SC_6_USPD16084632_HJ3FYBBXX_L2_2.fq.gz

Genome browser session
(e.g. UCSC)

Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.

Methodology

Replicates

Two parental hiPSCs lines (III-3; WT/MT, IV-1; WT/WT), and two isogeneic lines (III-3; WT/Corr-WT, IV-1; WT/Ins-MT).
Two different antibodies were used for LMNA ChIP-seq. ab8984 and sc-376248

Sequencing depth

Raw_Reads Clean_Reads Raw_Base(G) Clean_Base(G) Effective_Rate(%) Error_Rate(%) Q20(%) Q30(%) GC_Content(%) Paired-End
H3K4_pt3WT 54856219 42145611 16.5 12.6 76.83 0.02 95.06 88.98 56.01 Yes
H3K4_pt3MT 59207590 39834003 17.8 12 67.28 0.01 95.73 90.23 57.05 Yes
H3K27_pt3WT 41638326 31134930 12.5 9.3 74.77 0.01 95.98 90.68 46.75 Yes
H3K27_pt3MT 53964775 40732993 16.2 12.2 75.48 0.01 96.2 91.15 47.07 Yes
ab_pt3WT 53766870 43207940 16.1 13 80.36 0.02 95.93 90.19 40.85 Yes
ab_pt3MT 47369372 41177310 14.2 12.4 86.93 0.02 95.56 89.41 41.44 Yes
SC_pt3WT 57419745 51251376 17.2 15.4 89.26 0.02 95.63 89.53 40.58 Yes
SC_pt3MT 50700560 48187024 15.2 14.5 95.04 0.02 95.45 89.2 40.92 Yes

Antibodies

Santa Cruz Biotechnology sc-2025: IgG
Active Motif 39159: H3K4me3
Active Motif 39155: H3K27me3
Abcam ab8984: LMNA
Santa Cruz Biotechnology sc-376248: LMNA

Peak calling parameters

Enriched Domain Detector (v1.0) with an 11 Kb bin size, gap penalty of 5, and FDR significance threshold of 0.05.

Data quality

LMNA-sc-376248: EDD peaks (pt3WT; 587, pt3MT; 614)
LMNA-ab8974: EDD peaks (pt3WT; 585, pt3MT; 615)

Software

- [ChIP-seq] AQUAS pipeline from the Kundaje lab at Stanford University (https://github.com/kundajelab/chip-seq-pipeline2)
- [ChIP-seq] duplicate reads were removed: MarkDuplicates from Picard Tools (v2.17.3)
- [ChIP-seq] LMNA data were analyzed: Enriched Domain Detector (v1.0) with an 11 Kb bin size, gap penalty of 5, and FDR significance threshold of 0.05.
- [ChIP-seq] LAD gain, loss, and intersection were found: bedtools (v2.27.1)

ARTICLE

https://doi.org/10.1038/s41586-019-1449-z

Plant cell-surface GIPC sphingolipids sense salt to trigger Ca^{2+} influx

Zhonghao Jiang[1,2,3,4], Xiaoping Zhou[3,8], Ming Tao[1,8], Fang Yuan[2,3,8], Lulu Liu[2,3,8], Feihua Wu[1,2,3,8], Xiaomei Wu[3], Yun Xiang[2], Yue Niu[2], Feng Liu[2], Chijun Li[2], Rui Ye[2], Benjamin Byeon[2], Yan Xue[2], Hongyan Zhao[3], Hsin-Neng Wang[4,5], Bridget M. Crawford[4,5], Douglas M. Johnson[6], Chanxing Hu[2], Christopher Pei[2], Wenming Zhou[1], Gary B. Swift[6], Han Zhang[7], Tuan Vo-Dinh[4,5], Zhangli Hu[1]*, James N. Siedow[2] & Zhen-Ming Pei[2]*

Salinity is detrimental to plant growth, crop production and food security worldwide. Excess salt triggers increases in cytosolic Ca^{2+} concentration, which activate Ca^{2+}-binding proteins and upregulate the Na^+/H^+ antiporter in order to remove Na^+. Salt-induced increases in Ca^{2+} have long been thought to be involved in the detection of salt stress, but the molecular components of the sensing machinery remain unknown. Here, using Ca^{2+}-imaging-based forward genetic screens, we isolated the *Arabidopsis thaliana* mutant *monocation-induced* $[Ca^{2+}]_i$ *increases 1* (*moca1*), and identified MOCA1 as a glucuronosyltransferase for glycosyl inositol phosphorylceramide (GIPC) sphingolipids in the plasma membrane. MOCA1 is required for salt-induced depolarization of the cell-surface potential, Ca^{2+} spikes and waves, Na^+/H^+ antiporter activation, and regulation of growth. Na^+ binds to GIPCs to gate Ca^{2+} influx channels. This salt-sensing mechanism might imply that plasma-membrane lipids are involved in adaption to various environmental salt levels, and could be used to improve salt resistance in crops.

More than 6% of the world's total land area and about 20% of irrigated land (which produces one-third of the world's food) are increasingly affected by salt buildup[1]. Excessive salt is detrimental to plant growth and development, and causes agricultural loss and severe deterioration of plant ecosystems[1,2]. Sodium chloride is the most soluble and widespread salt found in soils. Sodium is not an essential nutrient in plants, and plants have evolved mechanisms to reduce intracellular sodium buildup[1,3]. In plants, high salinity triggers early short-term responses for perceiving and transducing the stress signal, and subsequent long-term responses for remodelling the transcriptional network to regulate growth and development. Although several molecular components in the early signalling pathway have been identified, plant salt sensors remain unknown[3–8].

Salt stress triggers increases in cytosolic free Ca^{2+} concentration ($[Ca^{2+}]_i$)[9,10], and the expulsion of excess intracellular Na^+ involves the Ca^{2+}-related salt-overly-sensitive (SOS) pathway[3,5]. The SOS pathway comprises the Ca^{2+} sensor SOS3 (a calcineurin B-like protein (also known as CBL4)), the protein kinase SOS2 (also known as CIPK24), and the Na^+/H^+ antiporter SOS1. Although salt-induced increases in $[Ca^{2+}]_i$ are thought to act as a detection mechanism, the molecular components involved in these increases are unknown[3–8,11]. In animals, sodium is an essential nutrient, and dedicated mechanisms have evolved to detect attractive low salt and aversive high salt conditions[12]. Notably, several ion channels act as salt-sensing taste receptors[13–16]. Sodium also triggers $[Ca^{2+}]_i$ spikes that are mediated by these salt-sensing channels. However, homologues of these channels do not exist in sequenced plant genomes.

High salinity increases both osmotic pressure and ionic strength, so salt can exert two stress effects: osmotic and ionic[1,4]. Ca^{2+}-imaging-based forward genetic screens have previously been used to isolate *Arabidopsis* mutants defective specifically in osmotic stress-induced Ca^{2+} increases, resulting in cloning of the osmosensing OSCA1 Ca^{2+} channel[17]. Here we have optimized experimental conditions for similar Ca^{2+}-imaging-based genetic screens to distinguish the ionic effect from the osmotic effect of salt stress. In this way, we isolated *Arabidopsis* mutants defective specifically in ionic stress-induced increases in $[Ca^{2+}]_i$. Analysis of a mutant identified through these screens revealed that plant-specific GIPC sphingolipids are involved in sensing salt-associated ionic stress in the plasma membrane.

moca1 is defective in salt-induced Ca^{2+} spikes

We attempted to identify ion-specific sensing mechanisms by using the same genetic approaches that were used to identify the osmosensing *osca1* mutant[17]. First, we needed to establish conditions under which the ionic effect of NaCl on $[Ca^{2+}]_i$ elevation was large whereas its effect on osmotic $[Ca^{2+}]_i$ elevation was minimal. We analysed the dose-dependent $[Ca^{2+}]_i$ increases induced by NaCl (ionic + osmotic effects) and sorbitol (osmotic effects only) using aequorin-based Ca^{2+} imaging. Throughout the range of concentrations tested, NaCl was more potent in triggering $[Ca^{2+}]_i$ increases than sorbitol at a similar osmolality (Fig. 1a; Extended Data Fig. 1). We reasoned that a threshold of 200 mM NaCl, at which the ionic effect was the highest and the osmotic effect was negligible, could be used to screen for mutants impaired in increases in $[Ca^{2+}]_i$ induced by ionic but not osmotic stresses.

Because it was difficult to physically map ethyl methanesulfonate (EMS)-induced mutations for the identification of aequorin-expressing *osca1*[17], we generated aequorin-expressing *Arabidopsis* populations mutagenized by transfer DNA (T-DNA) insertions using the vector pBIB-BASTA. We screened around 86,000 T2 seeds, and recovered about 10,000 seedlings in which increases in $[Ca^{2+}]_i$ in response to 200 mM NaCl were low. These lines were retested individually for four generations, and six individual lines with stable phenotypes were isolated as putative mutants. We then analysed several phenotypes to prioritize further characterization: 1) the plants did not have apparent

[1]College of Life Sciences and Oceanography, Longhua Innovation Institute for Biotechnology, Shenzhen University, Shenzhen, China. [2]Department of Biology, Duke University, Durham, NC, USA. [3]College of Life and Environmental Sciences, Hangzhou Normal University, Hangzhou, China. [4]Fitzpatrick Institute for Photonics, Duke University, Durham, NC, USA. [5]Department of Biomedical Engineering, Duke University, Durham, NC, USA. [6]Department of Physics, Duke University, Durham, NC, USA. [7]College of Optoelectronic Engineering, Shenzhen University, Shenzhen, China. [8]These authors contributed equally: Xiaoping Zhou, Ming Tao, Fang Yuan, Lulu Liu, Feihua Wu. *e-mail: huzl@szu.edu.cn; zpei@duke.edu

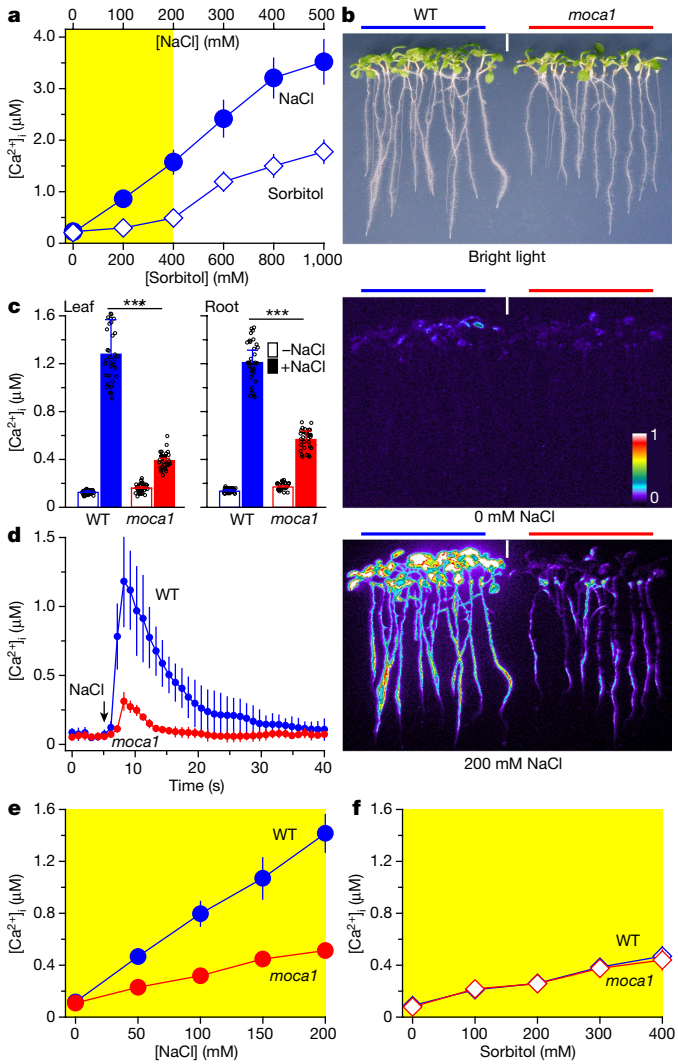


Fig. 1 | Isolation of *moca1* mutant defective in NaCl-induced increases in $[Ca^{2+}]_i$. **a**, Elevation in $[Ca^{2+}]_i$ in wild-type (WT) plants expressing aequorin plotted as a function of applied concentrations of NaCl and sorbitol. Data from four representative experiments are shown (mean ± s.d.; $n = 32$ seedlings; two-way ANOVA, $P < 0.001$). **b**, Aequorin imaging of NaCl-induced increases in $[Ca^{2+}]_i$ in plants treated with water or 200 mM NaCl. $[Ca^{2+}]_i$ is shown on a pseudo-colour scale (middle, bottom). Similar results were seen in more than 50 independent experiments. **c**, Quantification of increases in $[Ca^{2+}]_i$ in leaves and roots from experiments similar to those in **b**. Data from five representative experiments are shown (mean ± s.d.; $n = 32$ seedlings; $^{***}P < 0.001$). **d**, Time-course analysis of NaCl-induced increases in $[Ca^{2+}]_i$. Seedlings were treated with 200 mM NaCl, and bioluminescence was recorded at intervals of 1 s. Data from four representative experiments are shown (mean ± s.d.; $n = 18$ seedlings; two-way ANOVA, $P < 0.001$). **e**, Averaged increases in $[Ca^{2+}]_i$ plotted as a function of applied [NaCl]. Data from four separate experiments are shown (mean ± s.d.; $n = 32$ seedlings; two-way ANOVA, $P < 0.001$). **f**, Increases in $[Ca^{2+}]_i$ plotted as a function of applied [sorbitol] (mean ± s.d.; $n = 32$; two-way ANOVA, $P = 0.079$).

growth and developmental phenotypes; 2) the Ca^{2+} phenotype was confirmed using another Ca^{2+} indicator; 3) the Ca^{2+} phenotype was shown to be specific to ionic stress; 4) physiological responses to salt stress were compromised; 5) the biochemical function of the mutated gene was directly related to Ca^{2+}-associated salt sensing. Here, we report on the most affected mutant, *monocation-induced $[Ca^{2+}]_i$ increases 1* (*moca1*).

The *moca1* seedlings did not display morphological, growth, or developmental phenotypes throughout their lifecycle (Extended Data Fig. 2a–g; Fig. 1b). Changes in $[Ca^{2+}]_i$ in response to water treatment were similar in *moca1* and wild-type plants, whereas *moca1* plants showed much lower levels of $[Ca^{2+}]_i$ in response to treatment with 200 mM NaCl than did wild-type plants, in both leaves and roots (Fig. 1b, c; Extended Data Fig. 2h, i). We analysed the kinetics of $[Ca^{2+}]_i$ elevation and detected lower peaks in *moca1* plants (Fig. 1d). Increases in $[Ca^{2+}]_i$ induced by concentrations of NaCl up to 200 mM were lower in *moca1* (Fig. 1e; Extended Data Fig. 2j, k), but increases in $[Ca^{2+}]_i$ induced by sorbitol up to 400 mM were similar in *moca1* and wild-type plants (Fig. 1f; Extended Data Fig. 2l, m).

We confirmed the *moca1* phenotype using the GFP-based yellow cameleon 3.6 (YC3.6) Ca^{2+} indicator[7,17,18]; NaCl-induced increases in $[Ca^{2+}]_i$ were lower in *moca1* roots (Fig. 2a, b). To test whether *moca1* and *osca1* are specific to ionic and osmotic sensing, respectively, we observed that with respect to NaCl-induced increases in $[Ca^{2+}]_i$, *osca1* and wild-type seedlings grown side-by-side were similar; whereas with respect to sorbitol-induced increases in $[Ca^{2+}]_i$, *moca1* and wild-type seedlings were similar (Extended Data Fig. 3a–e), showing that *moca1* and *osca1* differ and that *moca1* has ionic-specific effects.

The *moca1* mutant is hypersensitive to salt stress

Salt treatment inhibited both leaf and root growth in *moca1* plants (Fig. 2c, d). Under mild salt stress of 60 mM NaCl and low medium Ca^{2+} concentrations, the survival rate of *moca1* plants was reduced (Fig. 2c, e). We investigated whether *moca1* affects the SOS pathway; that is, whether salt-induced increases in $[Ca^{2+}]_i$ are necessary and sufficient for triggering the SOS3 → SOS2 → SOS1 signalling relay[5,7,8]. Although Na^+/H^+ antiporter activity was similar in *moca1* and wild-type plants, treatment with NaCl enhanced the activity in wild-type but not *moca1* plants (Fig. 2f; Extended Data Fig. 3f). As the concentration of NaCl in the medium increased, the Na content in *moca1* plants increased more, whereas the K content decreased more than these in wild type (Fig. 2g, h), consistent with the mutant's hypersensitivity to salt and reduced Na^+/H^+ antiporter activity. These results fill the gap in Ca^{2+} elevation between NaCl stress and the SOS pathway[5,7], showing that *moca1* is upstream of SOS. Nevertheless, seedling growth in response to osmotic stress and abscisic acid was not different from the wild type in *moca1* plants (Extended Data Fig. 4a–d), revealing that *moca1* does not have pleiotropic defects in stress response.

moca1 lacks cation-evoked Ca^{2+} spikes and waves

To determine whether *moca1* is specific to ionic stress as compared to other stimuli known to trigger increases in $[Ca^{2+}]_i$[5,7,8,19], we measured increases in $[Ca^{2+}]_i$ in response to H_2O_2, cold temperature, and high external Ca^{2+}. These responses were nearly identical in *moca1* and wild-type plants (Fig. 3a–c; Extended Data Fig. 5a–f). Increases in $[Ca^{2+}]_i$ in response to other monovalent cations, such as K^+ and Li^+, were reduced in *moca1* plants, and the substitution of Cl^- for NO_3^- did not affect the *moca1* Ca^{2+} phenotype (Fig. 3d–f; Extended Data Fig. 5g–l), showing that *moca1* is mainly selective for monovalent cations. Note that KCl and LiCl stresses inhibited the growth of wild-type and *moca1* plants to the same degree (Extended Data Fig. 4e–h), suggesting that although the initial Ca^{2+}–SOS signalling mechanism could be identical for Na^+, K^+, and Li^+, the selectivity of SOS1 for Na^+ over K^+ and Li^+ might explain the attenuated growth of *moca1* only under Na^+ stress.

In plants, long-distance signalling—such as plasma membrane potential depolarization and $[Ca^{2+}]_i$ waves—has been proposed to be a response to environmental stresses, including salt stress[18,20]. Localized salt stress at the root tip causes increases in $[Ca^{2+}]_i$, which initiate a $[Ca^{2+}]_i$ wave that propagates to distal shoot tissues[21,22]. Using YC3.6 fluorescence resonance energy transfer (FRET)-based Ca^{2+} imaging, we observed that the NaCl-triggered $[Ca^{2+}]_i$ wave was almost completely absent in *moca1* roots (Fig. 3g–i; Supplementary Video 1). As expected, the Ca^{2+} channel inhibitor La^{3+} blocked the $[Ca^{2+}]_i$ wave in wild-type roots (Extended Data Fig. 6; $P < 0.01$), confirming that Ca^{2+} signals are propagated from the root tip[21,22]. Previous studies have shown that the speed but not the initiation of the Ca^{2+} wave is altered in the *tpc1* mutant[21,22], placing TPC1 downstream of MOCA1.

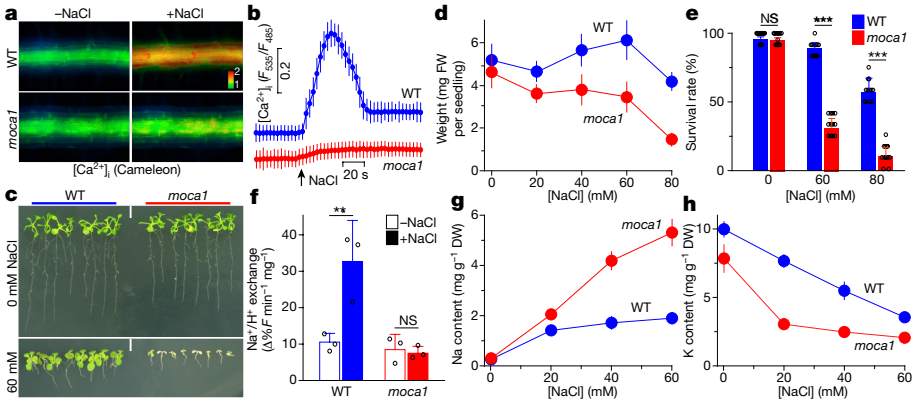


Fig. 2 | The *moca1* mutant is defective in the SOS pathway and hypersensitive to salt stress. a, **b**, Increases in $[Ca^{2+}]_i$ induced by NaCl in roots. YC3.6 emission images were taken every 3 s, and 200 mM NaCl was added at the time indicated (**a**). Emission ratios are shown using a pseudo-colour scale and quantified from experiments similar to these in **a** (**b**; mean ± s.d.; $n = 10$ seedlings). Similar results were seen in more than ten independent experiments. **c**, Plants were grown on half-strength Murashige and Skoog (½ MS) medium containing 0.2 mM $CaCl_2$ with or without 60 mM NaCl for 12 days. Similar results were seen in more than ten independent experiments. **d**, **e**, Fresh weight (FW; **d**) and survival rate (**e**) from experiments similar to those in **c** were quantified. Data are from five independent experiments (mean ± s.d.; $n = 12$ pools (8–12 seedlings per pool); two-way ANOVA, $P < 0.001$; NS, not significant; $^{***}P < 0.001$). **f**, Na^+/H^+ exchange activity from plants treated with water or 100 mM NaCl for 24 h (mean ± s.d.; $n = 3$; $^{**}P < 0.01$). **g**, **h**, The content of Na (**g**) and K (**h**) of plants from experiments similar to those in **c**. Data are presented as mean ± s.e.m. ($n = 6$; two-way ANOVA, $P < 0.001$).

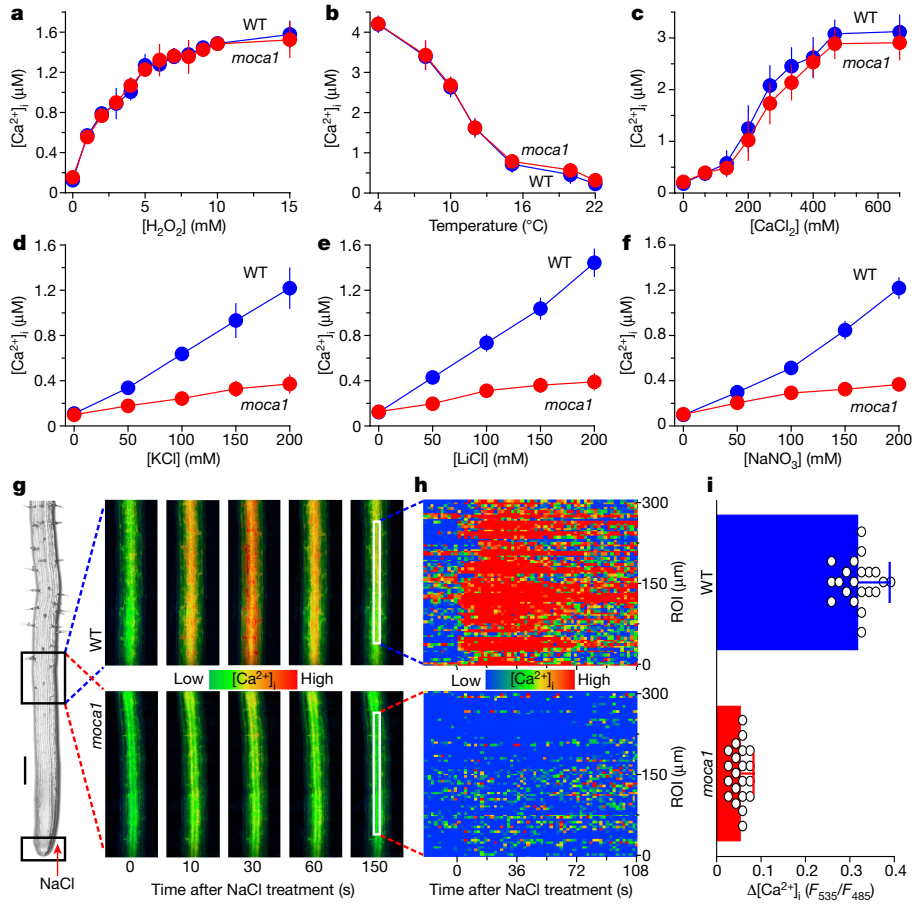


Fig. 3 | The *moca1* mutant abolishes $[Ca^{2+}]_i$ spikes and waves induced by monovalent cations. a–**c**, Increases in $[Ca^{2+}]_i$ in seedlings plotted as a function of applied $[H_2O_2]$ (**a**), temperature gradient (**b**), and $[CaCl_2]$ (**c**). Data are from three independent experiments similar to those in Fig. 1a–c (mean ± s.d.; $n = 32$ seedlings; two-way ANOVA; **a**, $P = 0.64$; **b**, $P = 0.06$; **c**, $P = 0.11$). **d**–**f**, Increases in $[Ca^{2+}]_i$ plotted as a function of applied [KCl] (**d**), [LiCl] (**e**), and $[NaNO_3]$ (**f**). Data are from three independent experiments (mean ± s.d.; $n = 32$ seedlings; two-way ANOVA, $P < 0.001$). **g**–**i**, Wave-like propagation of $[Ca^{2+}]_i$ through the root 1,000 μm from the site at which root-tip was treated with 200 mM NaCl was analysed via YC3.6 imaging. Images in regions of interest (ROIs) at indicated times are shown (**g**; scale bar, 200 μm). The wave dynamics of $[Ca^{2+}]_i$ were examined shootward from roots similar to the ROIs in **g** by repeatedly analysing images taken every 2 s, and data from five roots were averaged and pseudo-colour coded (**h**). Similar results were seen in more than 20 independent experiments. Changes in peaks of $[Ca^{2+}]_i$ waves in ROIs were quantified from experiments similar to those in **g** (**i**; mean ± s.d.; $n = 21$ roots; Student's *t*-test, $P < 0.001$).

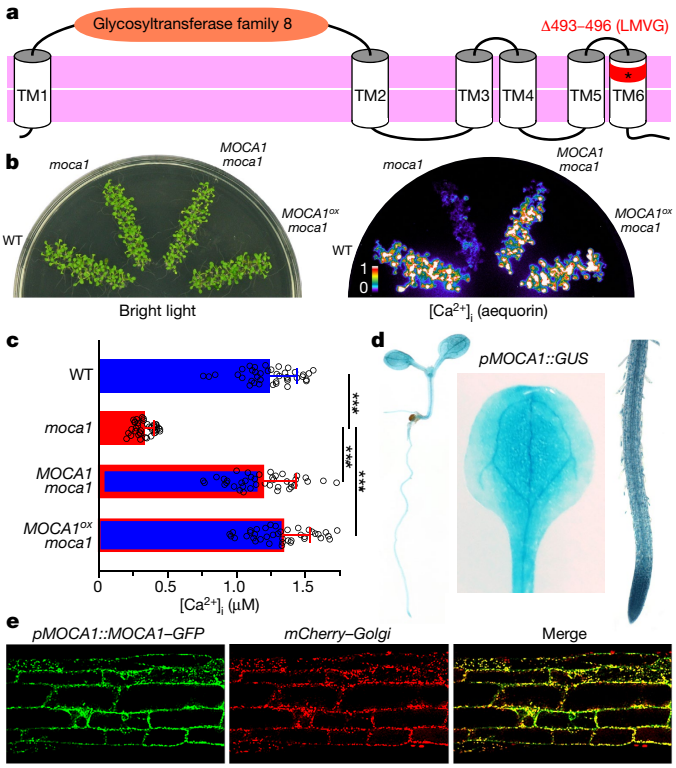


Fig. 4 | MOCA1 encodes a glucuronosyltransferase. a, The predicted topology of MOCA1. TM, transmembrane domain; asterisk, Δ493–496 (LMVG), the four amino acid residues that are deleted in the mutant protein product of *moca1*. **b**, **c**, Complementation of *moca1* Ca^{2+} phenotype by expressing MOCA1 driven by its own promoter (*MOCA1 moca1*) or 35S promoter (*MOCA1^{ox} moca1*). Aequorin images of seedlings treated with 200 mM NaCl are shown (**b**), and increases in $[Ca^{2+}]_i$ resulting from experiments as in **b** are quantified (**c**; mean ± s.d.; $n = 40$ pools (30 seedlings per pool); Student's *t*-test, $^{***}P < 0.001$). **d**, Expression patterns of *pMOCA1::GUS* in the seedling, leaf, and root. Similar results were seen in more than ten independent experiments. **e**, Golgi membrane localization of MOCA1 in root epidermal cells co-expressing MOCA1 promoter-driven MOCA1–GFP (*pMOCA1::MOCA1–GFP*) and a Golgi marker tagged with mCherry (*mCherry–Golgi*). Similar results were seen in more than ten independent experiments.

Together, our results demonstrate that *moca1* mutants are severely defective in major early salt signalling events ($[Ca^{2+}]_i$ spikes or waves and SOS1 activation) as well as showing attenuated growth and development in response to salt stress, implying that salt detection might be disrupted in this mutant.

MOCA1 encodes a glucuronosyltransferase for GIPCs

As the T-DNA insertion in *moca1* was lost, we carried out positional mapping[17]. Segregation analysis showed that the *moca1* phenotype was caused by a recessive mutation in a single nuclear gene (Extended Data Fig. 7a). We crossed *moca1* to the Wassilewskija ecotype to generate mapping lines, and mapped *moca1* to the upper arm of chromosome 5 (Extended Data Fig. 7b). Through fine mapping and candidate gene sequencing, we identified a 12-nucleotide deletion in a gene (*At5g18480*; Extended Data Fig. 7c, d; Fig. 4a), that encodes inositol phosphorylceramide glucuronosyltransferase 1 (IPUT1)[23]. IPUT1 is classified as glycogenin-like starch initiation protein 6 (PGSIP6) in the glucuronosyltransferase subfamily 8 (GT8)[24]. Hydrophobicity analyses predicted that MOCA1 is an integral membrane protein with six transmembrane (TM) α-helices and a long stretched GT8 domain between TM1 and TM2, and the *moca1* mutant has a four-amino-acid-residue deletion in TM6 (Extended Data Fig. 7d, e; Fig. 4a).

To confirm that *MOCA1* is the relevant gene, we found that expression of *MOCA1* from the endogenous promoter (*MOCA1 moca1*) or overexpression driven by the 35S promoter (*MOCA1^{ox} moca1*) could complement the *moca1* Ca^{2+} phenotype (Fig. 4b, c; Extended Data Fig. 8a, b). The salt hypersensitivity phenotypes in root growth and viability were also complemented to wild-type levels (Extended Data Fig. 8c–e; $P > 0.05$). We tried to generate homozygous *moca1-2* and *moca1-3* T-DNA insertion mutants, but these mutations were lethal (as seen in *iput1-1* and *iput1-2* mutants[23]). Note that *moca1* is *moca1-1* in this study.

Analysis of transgenic plants expressing a GUS (β-glucuronidase) reporter gene driven by the *MOCA1* promoter (*pMOCA1::GUS*) showed that MOCA1 is broadly expressed, particularly in cotyledons and roots, consistent with the tissues where Ca^{2+} phenotypes and physiological salt sensitivity were observed (Fig. 4d; Extended Data Fig. 8f). We analysed fluorescence in *pMOCA1::MOCA1–GFP* lines that contained a Golgi marker (*mCherry–Golgi*) and observed punctate patterns co-localized with the Golgi marker in the cytosol in root epidermal cells (Fig. 4e), consistent with PGSIP6 (MOCA1)–GFP localization and Golgi proteome analyses[24,25]. The four-amino-acid-residue deletion did not affect the subcellular localization of mutant MOCA1 (mMOCA1) (Extended Data Fig. 8g). The remaining question was how MOCA1 governs salt-induced increases in $[Ca^{2+}]_i$.

Na^+ ions bind to GIPC sphingolipids

IPUT1 transfers a glucuronic acid (GlcA) residue from UDP–GlcA to inositol phosphorylceramide (IPC) to form GIPCs[23] (Extended Data Fig. 8h). The *iput1* mutant rescued by expressing pollen-specific promoter-driven IPUT1 contains low levels of GIPCs and is a severe dwarf[26]. We measured GIPCs and IPCs, and found that *moca1* plants contained lower levels of GIPCs but higher levels of IPCs than the wild type (Fig. 5a–d). *moca1* and wild-type plants grown in agar plates and soil were indistinguishable throughout their life cycle (Figs. 1b, 2c, 4b; Extended Data Fig. 2a–g), in contrast to the severe phenotypes seen in *iput1* and *iput1*-rescued lines[23,26]. Thus, the GIPC levels might be above the threshold required for normal growth and development in *moca1* plants, while being low enough to compromise salt sensing.

GIPCs are a major class of lipids in fungi, protozoans, and plants, but not in animals, and are abundant in the plasma membrane (about 25% of total lipids)[27–30]. GIPCs have very long saturated acyl chains; and have been proposed to be located in the outer leaflet of the plasma membrane and enriched in raft-like lipid micro-domains[27,29–31]. Although several hypotheses have been proposed regarding the role of GIPCs, including cell wall anchoring, lipid moieties for protein anchoring, cell-surface recognition, and precursors of signalling molecules, their exact roles remain unclear[27,28]. Conversely, in animals, sphingolipids play a central role in cell signalling[29,31], and perturbations of sphingolipids lead to various human diseases[32]. The negatively charged GIPCs are structural homologues of animal gangliosides[27,28], which regulate receptors and ion channels as well as Ca^{2+} homeostasis[31,32].

On the basis of this information, we hypothesized that the negatively charged GIPCs could provide Na^+-binding sites on the cell surface and gate Ca^{2+} influx channels in plants, as seen in the regulation of channels in animals[33]. We first measured changes in cell-surface potentials in response to Na^+ treatment. In wild-type protoplasts, increases in NaCl concentration led to increases in the ζ potentials, which largely represent cell-surface potentials (Fig. 5e). In *moca1* protoplasts, the ζ potentials did not respond to NaCl treatment, and were decreased slightly (Fig. 5e). In the absence of NaCl, ζ potentials were consistently lower in *moca1* protoplasts than in wild-type protoplasts, further demonstrating the marked alteration in electric charges in the plasma membrane. Treatment with up to 15 mM NaCl did not significantly affect the integrity of wild-type protoplasts (survival rate ~85%; Fig. 5f), but lowered *moca1* survival rates to about 25%. Without NaCl treatment, the survival rates for wild-type and *moca1* protoplasts were about 85% and 65%, respectively (Fig. 5f), showing that *moca1* protoplasts were hypersensitive to salt stress. These results suggested that GIPCs might directly detect Na^+ levels in the apoplastic space.

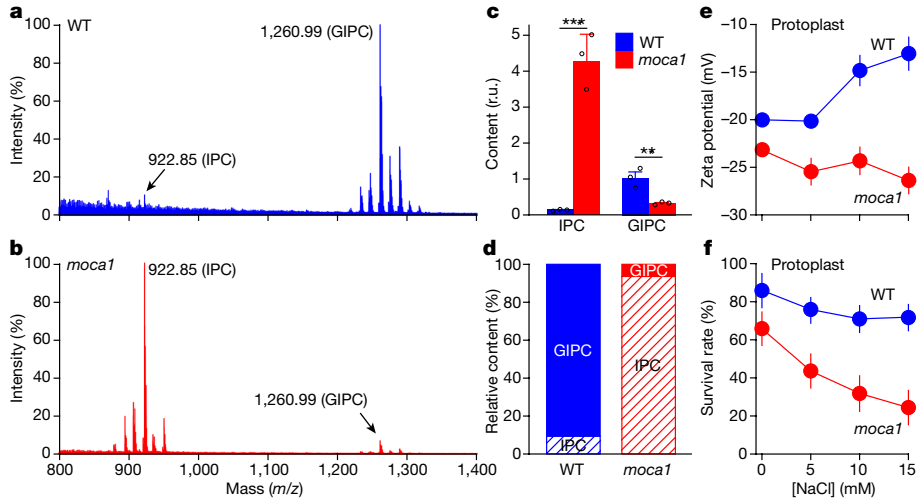


Fig. 5 | MOCA1-related GIPCs are responsible for NaCl-induced changes in cell-surface potentials. a, **b**, Matrix-assisted laser desorption/ionization-mass spectrometry (MALDI-MS) analysis of GIPCs extracted from wild-type (**a**) or *moca1* leaves (**b**). IPC and GIPC series A differ by their number of saccharide units, from zero (IPC) to two (GIPC series A). **c**, **d**, Relative contents of IPC and GIPC series A in leaves from experiments as in **a** and **b** were quantified with wild-type GIPC content normalized as 1 (r.u., relative unit; **c**), or as a relative percentage between IPC and GIPC with the total of IPC and GIPC as 100% (**d**). Data are presented as mean ± s.d. ($n = 3$; Student's *t*-test, $^{***}P < 0.001$, $^{**}P < 0.01$). **e**, ζ potentials of mesophyll cell protoplasts plotted as a function of applied [NaCl]. Date are from four representative experiments (mean ± s.d.; $n = 20$ protoplasts per data point; two-way ANOVA, $P < 0.001$). **f**, Survival rates of protoplasts plotted as a function of applied [NaCl]. Data are presented as mean ± s.d. ($n = 5$ repeats (15 protoplasts); two-way ANOVA, $P < 0.001$).

We investigated whether Na^+ ions bind to GIPCs using isothermal titration calorimetry (ITC). We produced lipid vesicles from the mass-spectrometry-analysed GIPC–IPC mixtures extracted from wild-type and *moca1* seedlings (Fig. 5a–d). The mixtures contained more than 90% GIPCs and more than 90% IPCs for wild-type and *moca1* seedlings, respectively (Fig. 5d). Thus, the properties of Na^+ binding to these lipids using wild-type or *moca1* mixtures approximate the properties of Na^+ binding to GIPCs or IPCs, respectively. ITC showed that Na^+ binds to GIPCs and IPCs with a dissociation constant (K_d) of 0.315 ± 0.083 mM and 0.286 ± 0.063 mM, respectively (Fig. 6a, b; Extended Data Fig. 9a–f; $P > 0.05$). The number of apparent binding sites for GIPCs and IPCs was 1.09 ± 0.02 and 0.77 ± 0.09, respectively (Extended Data Fig. 9c; $P < 0.01$), in agreement with the presence of negative charges on PO_4^- and $GlcA^-$ in GIPCs and only PO_4^- in IPCs. Other thermal properties of GIPCs and IPCs were similar (Extended Data Fig. 9d–f). We also analysed the binding ability of K^+ and Li^+ to GIPCs and IPCs (Extended Data Fig. 9g–j). For all three ions, there were consistently more binding sites in GIPCs than in IPCs, whereas for a given cation the binding affinities to GIPCs and IPCs were similar.

Given that lipid micro-domains in the plant plasma membrane contain a large amount of GIPCs[27,30], these results indicate that disruption of GIPC biosynthesis might reduce GIPC content in *moca1*, leading to the reduction in Na^+-binding sites and thereby preventing the subsequent cell-surface potential depolarization that gates Ca^{2+} channels (Fig. 6c).

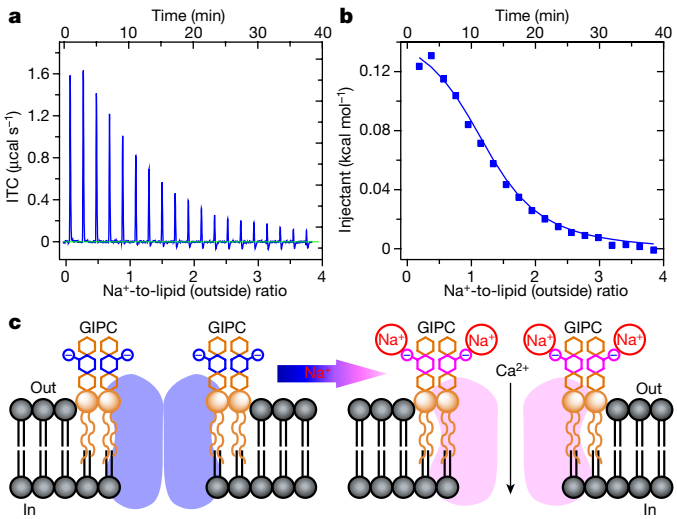


Fig. 6 | Na^+ binds to GIPCs and gates Ca^{2+} influx channels. a, **b**, ITC analysis of Na^+ binding to GIPCs from wild-type GIPC–IPC mixture with >90% GIPCs. ITC data (**a**) and plots of injected heat for NaCl injections into the sample cell are shown (**b**). Six independent experiments were performed, and similar results were obtained. **c**, Model of how GIPCs sense salt and gate Ca^{2+} influx channels.

Discussion

We have tackled the long-standing issue of whether salt-induced Ca^{2+} signalling serves as a salt-sensing mechanism in plants. We identified the first molecular component required for salt-induced $[Ca^{2+}]_i$ elevation, MOCA1, and revealed a biochemical function of GIPCs as monovalent-cation sensors. Notably, considering that the GT8 family members in non-plant organisms do not have multiple transmembrane domains, the mutation in TM6 in *moca1* may suggest that in plants, unique multiple transmembrane domains in the GT8 family have evolved as an adaptation to salt stress environments (Extended Data Fig. 10). Furthermore, identification of GIPCs as sensors in the salt CBL–CIPK pathway will shed light on the molecular mechanisms that underlie the detection of other nutrients via CBL–CIPK pathways[7,8,34–36].

We propose a working model for plant salt cation sensing (Fig. 6c): Na^+ ions bind to GIPCs, and depolarize the cell-surface potential to gate Ca^{2+} influx channels. The functioning of ion channels and receptors in the membrane depends critically on how their transmembrane segments are embedded in the membrane[29,31,32], and the regulation of ion channels by cell-surface potentials was recorded more than 40 years ago in animals[33], although the exact molecular mechanisms of this process remain unknown. Sphingolipids are structural components of membranes found in lipid micro-domains, and also act as intracellular second messengers in animals[30,32–34]. However, not much is known about the binding of sphingolipids to ion channels to gate them via cell-surface charges. On the other hand, phosphatidylinositol binds ion channels in the cytoplasmic leaflet and regulates ion channel function[37–40]. Evidently, GIPC-mediated salt sensing does not resemble any known sensory system found in other organisms. Our findings allow us to propose that rather than a sole salt sensor, cation-sensing

GIPCs, osmosensing OSCA1 and components yet to be defined may work together to integrate both ionic and osmotic aspects of salt into the salt Ca^{2+} signalling pathway in plants.

Much progress has been made over the last three decades in understanding the phenomenon of $[Ca^{2+}]_i$ elevation in response to abiotic and biotic stimuli in plants[5,7,41,42]. Ca^{2+}-imaging-based genetic screens have led to the identification of only a few receptors or sensors, including DORN1 for external ATP, OSCA1 for osmotic stress, and LORE for lipopolysaccharides[17,43,44]. These Ca^{2+}-related receptors could be classified into three groups: (1) receptor-like kinases, including NFR1, NFP (also known as NFR5) and DMI2 (also known as SYMRK) for Nod factors, FLS2 for flg22, EFR for elf18 and elf26, PEPR1 for AtPep1, and FER for rapid alkalinization factor, as well as DORN1 and LORE; (2) the receptor channel OSCA1; and (3) transmembrane receptors[8,41]. Ca^{2+} channels that are not receptors or sensors but are responsible for Ca^{2+} increases have also been found, such as DMI1, Pollux and Castor and CNGC15 for Nod factors, CNGC14 for auxin and CNGC18 for the pollen tube[41,45,46], and GLRs for wounding and sperm chemotaxis[47,48]. It is plausible to speculate that GIPC-associated Ca^{2+} channels belong to this category. In animals, Ca^{2+}-related receptors comprise G-protein coupled receptors, receptor tyrosine kinases, and receptor channels[12,49], and animal salt sensors are receptor channels[13–16]. Therefore, GIPC-mediated salt sensing in plants differs from all these receptors found in animals and plants. Note that GIPCs are receptors for pathogenic necrosis and ethylene-inducing peptide 1-like proteins in eudicot but not monocot plants[50], and gangliosides are receptors for axon–myelin interactions in animals[32].

In conclusion, our results shed light on salt sensing in plants, highlight the importance of GIPCs—as a specific class of sphingolipids—for the regulation (and modulation) of signalling processes at the plasma membrane, and underscore the functional versatility of various lipids in different evolutionary branches of life. Our findings could also provide potential molecular genetic targets for engineering salt-resistant crops.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1449-z.

Received: 25 April 2018; Accepted: 3 July 2019;
Published online 31 July 2019.

1. Munns, R. & Tester, M. Mechanisms of salinity tolerance. *Annu. Rev. Plant Biol.* **59**, 651–681 (2008).
2. Ismail, A. M. & Horie, T. Genomics, physiology, and molecular breeding approaches for improving salt tolerance. *Annu. Rev. Plant Biol.* **68**, 405–434 (2017).
3. Yang, Y. & Guo, Y. Elucidating the molecular mechanisms mediating plant salt-stress responses. *New Phytol.* **217**, 523–539 (2018).
4. Deinlein, U. et al. Plant salt-tolerance mechanisms. *Trends Plant Sci.* **19**, 371–379 (2014).
5. Zhu, J. K. Abiotic stress signaling and responses in plants. *Cell* **167**, 313–324 (2016).
6. Julkowska, M. M. & Testerink, C. Tuning plant signaling and growth to survive salt. *Trends Plant Sci.* **20**, 586–594 (2015).
7. Kudla, J. et al. Advances and current challenges in calcium signaling. *New Phytol.* **218**, 414–431 (2018).
8. Edel, K. H., Marchadier, E., Brownlee, C., Kudla, J. & Hetherington, A. M. The evolution of calcium-based signalling in plants. *Curr. Biol.* **27**, R667–R679 (2017).
9. Knight, M. R., Campbell, A. K., Smith, S. M. & Trewavas, A. J. Transgenic plant aequorin reports the effects of touch and cold-shock and elicitors on cytoplasmic calcium. *Nature* **352**, 524–526 (1991).
10. Knight, H., Trewavas, A. J. & Knight, M. R. Calcium signalling in *Arabidopsis thaliana* responding to drought and salinity. *Plant J.* **12**, 1067–1078 (1997).
11. Hedrich, R. Ion channels in plants. *Physiol. Rev.* **92**, 1777–1811 (2012).
12. Roper, S. D. & Chaudhari, N. Taste buds: cells, signals and synapses. *Nat. Rev. Neurosci.* **18**, 485–497 (2017).
13. Chandrashekar, J. et al. The cells and peripheral representation of sodium taste in mice. *Nature* **464**, 297–301 (2010).
14. Zhang, Y. V., Ni, J. & Montell, C. The molecular basis for attractive salt-taste coding in *Drosophila*. *Science* **340**, 1334–1338 (2013).
15. Oka, Y., Butnaru, M., Von Buchholtz, L., Ryba, N. J. P. & Zuker, C. S. High salt recruits aversive taste pathways. *Nature* **494**, 472–475 (2013).
16. Chatzigeorgiou, M., Bang, S., Hwang, S. W. & Schafer, W. R. *tmc-1* encodes a sodium-sensitive channel required for salt chemosensation in *C. elegans*. *Nature* **494**, 95–99 (2013).
17. Yuan, F. et al. OSCA1 mediates osmotic-stress-evoked Ca^{2+} increases vital for osmosensing in *Arabidopsis*. *Nature* **514**, 367–371 (2014).
18. Choi, W. G., Hilleary, R., Swanson, S. J., Kim, S. H. & Gilroy, S. Rapid, long-distance electrical and calcium signaling in plants. *Annu. Rev. Plant Biol.* **67**, 287–307 (2016).
19. Martí, M. C., Stancombe, M. A. & Webb, A. A. R. Cell- and stimulus type-specific intracellular free Ca^{2+} signals in *Arabidopsis*. *Plant Physiol.* **163**, 625–634 (2013).
20. Hedrich, R., Salvador-Recatalà, V. & Dreyer, I. Electrical wiring and long-distance plant communication. *Trends Plant Sci.* **21**, 376–387 (2016).
21. Choi, W. G., Toyota, M., Kim, S. H., Hilleary, R. & Gilroy, S. Salt stress-induced Ca^{2+} waves are associated with rapid, long-distance root-to-shoot signaling in plants. *Proc. Natl Acad. Sci. USA* **111**, 6497–6502 (2014).
22. Evans, M. J., Choi, W. G., Gilroy, S. & Morris, R. J. A ROS-assisted calcium wave dependent on the AtRBOHD NADPH oxidase and TPC1 cation channel propagates the systemic response to salt stress. *Plant Physiol.* **171**, 1771–1784 (2016).
23. Rennie, E. A. et al. Identification of a sphingolipid α-glucuronosyltransferase that is essential for pollen function in *Arabidopsis*. *Plant Cell* **26**, 3314–3325 (2014).
24. Rennie, E. A. et al. Three members of the *Arabidopsis* glycosyltransferase family 8 are xylan glucuronosyltransferases. *Plant Physiol.* **159**, 1408–1417 (2012).
25. Nikolovski, N., Shliaha, P. V., Gatto, L., Dupree, P. & Lilley, K. S. Label-free protein quantification for plant Golgi protein localization and abundance. *Plant Physiol.* **166**, 1033–1043 (2014).
26. Tartaglio, V. et al. Glycosylation of inositol phosphorylceramide sphingolipids is required for normal growth and reproduction in *Arabidopsis*. *Plant J.* **89**, 278–290 (2017).
27. Gronnier, J., Germain, V., Gouguet, P., Cacas, J.-L. & Mongrand, S. GIPC: glycosyl inositol phospho ceramides, the major sphingolipids on earth. *Plant Signal. Behav.* **11**, e1152438 (2016).
28. Markham, J. E., Lynch, D. V., Napier, J. A., Dunn, T. M. & Cahoon, E. B. Plant sphingolipids: function follows form. *Curr. Opin. Plant Biol.* **16**, 350–357 (2013).
29. Hannun, Y. A. & Obeid, L. M. Sphingolipids and their metabolism in physiology and disease. *Nat. Rev. Mol. Cell Biol.* **19**, 175–191 (2018).
30. Cacas, J. L. et al. Revisiting plant plasma membrane lipids in tobacco: a focus on sphingolipids. *Plant Physiol.* **170**, 367–384 (2016).
31. Ledeen, R. W. & Wu, G. The multi-tasked life of GM1 ganglioside, a true factotum of nature. *Trends Biochem. Sci.* **40**, 407–418 (2015).
32. Schnaar, R. L., Gerardy-Schahn, R. & Hildebrandt, H. Sialic acids in the brain: gangliosides and polysialic acid in nervous system development, stability, disease, and regeneration. *Physiol. Rev.* **94**, 461–518 (2014).
33. Green, W. N. & Andersen, O. S. Surface charges and ion channel function. *Annu. Rev. Physiol.* **53**, 341–359 (1991).
34. Liu, K. H. et al. Discovery of nitrate-CPK-NLP signalling in central nutrient-growth networks. *Nature* **545**, 311–316 (2017).
35. Ho, C. H., Lin, S. H., Hu, H. C. & Tsay, Y. F. CHL1 functions as a nitrate sensor in plants. *Cell* **138**, 1184–1194 (2009).
36. Xu, J. et al. A protein kinase, interacting with two calcineurin B-like proteins, regulates K^+ transporter AKT1 in *Arabidopsis*. *Cell* **125**, 1347–1360 (2006).
37. McLaughlin, S. & Murray, D. Plasma membrane phosphoinositide organization by protein electrostatics. *Nature* **438**, 605–611 (2005).
38. Balla, T. Phosphoinositides: tiny lipids with giant impact on cell regulation. *Physiol. Rev.* **93**, 1019–1137 (2013).
39. Hirschi, M. et al. Cryo-electron microscopy structure of the lysosomal calcium-permeable channel TRPML3. *Nature* **550**, 411–414 (2017).
40. Gao, Y., Cao, E., Julius, D. & Cheng, Y. TRPV1 structures in nanodiscs reveal mechanisms of ligand and lipid action. *Nature* **534**, 347–351 (2016).
41. Zipfel, C. & Oldroyd, G. E. D. Plant signalling in symbiosis and immunity. *Nature* **543**, 328–336 (2017).
42. Hamilton, E. S., Schlegel, A. M. & Haswell, E. S. United in diversity: mechanosensitive ion channels in plants. *Annu. Rev. Plant Biol.* **66**, 113–137 (2015).
43. Choi, J. et al. Identification of a plant receptor for extracellular ATP. *Science* **343**, 290–294 (2014).
44. Ranf, S. et al. A lectin S-domain receptor kinase mediates lipopolysaccharide sensing in *Arabidopsis thaliana*. *Nat. Immunol.* **16**, 426–433 (2015).
45. Charpentier, M. et al. Nuclear-localized cyclic nucleotide-gated channels mediate symbiotic calcium oscillations. *Science* **352**, 1102–1105 (2016).
46. Dindas, J. et al. AUX1-mediated root hair auxin influx governs $SCF^{TIR1/AFB}$-type Ca^{2+} signaling. *Nat. Commun.* **9**, 1174 (2018).
47. Mousavi, S. A. R., Chauvin, A., Pascaud, F., Kellenberger, S. & Farmer, E. E. *GLUTAMATE RECEPTOR-LIKE* genes mediate leaf-to-leaf wound signalling. *Nature* **500**, 422–426 (2013).
48. Ortiz-Ramírez, C. et al. GLUTAMATE RECEPTOR-LIKE channels are essential for chemotaxis and reproduction in mosses. *Nature* **549**, 91–95 (2017).
49. Murthy, S. E., Dubin, A. E. & Patapoutian, A. Piezos thrive under pressure: mechanically activated ion channels in health and disease. *Nat. Rev. Mol. Cell Biol.* **18**, 771–783 (2017).
50. Lenarčič, T. et al. Eudicot plant-specific sphingolipids determine host selectivity of microbial NLP cytolysins. *Science* **358**, 1431–1434 (2017).

METHODS

Plant material and growth conditions. *Arabidopsis thaliana* ecotype Col-0 and *Arabidopsis thaliana* Col-0 constitutively expressing the intracellular Ca^{2+} indicator aequorin (pMAEQUORIN2; a gift from M. Knight; Col-0 (aequorin))[9,10] and YC3.6 (a gift from S. Gilroy)[51] were used. Two *Arabidopsis* T-DNA insertion lines for At5g18480, SALK_131321 (*moca1-2*) and GK-856G03 (*moca1-3*), were obtained from the *Arabidopsis* Biological Resource Center (ABRC). Note that *moca1* is *moca1-1* in this study. *Arabidopsis* seedlings were grown in soil (Scotts Metro-Mix 200) or in Petri dishes in ½ MS (Sigma), 1.5% (w/v) sucrose (Sigma), and 0.6% (w/v) agar (Sigma) unless otherwise described in controlled environmental rooms or plant growth chambers (Percival Scientific) at 21 ± 2 °C. The fluency rate of white light was ~110 μmol photons m^{-2} s^{-1}. The photoperiods were 16 h light/8 h dark cycles. Seeds were sown on soil or MS medium, placed at 4 °C for 3 days in the dark, and then transferred to growth rooms or chambers.

Aequorin bioluminescence-based Ca^{2+} imaging. Cytosolic free Ca^{2+} concentration ($[Ca^{2+}]_i$) was measured using plants expressing aequorin as described previously[9,17]. *Arabidopsis* seedlings were applied evenly with 3.3 ml of 10 μM coelenterazine (Prolume) per 150 mm × 15-mm Petri dish 12 h before imaging and placed in the dark. Aequorin bioluminescence imaging was performed using either a cryogenically cooled and back-illuminated CCD camera ChemiPro HT system equipped with a light-tight box (Roper Scientific) or a newer version Lumazone system (Pylon1300B, Roper Scientific) equipped with the H-800 light-tight controlled environmental box (Bio-One Scientific Instrument). A liquid nitrogen autofiller (Roper Scientific or Bio-One Scientific Instrument) was attached to the imaging system to provide constant cooling. The cameras were controlled by WinView/32 (Roper) and bioluminescence images were analysed using MetaMorph 6.3 or MetaMorph Basic Acquisition for Microscope (Molecular Devices). The recording of luminescence (L) was started 10 s before treatments and lasted for 3 min unless otherwise described. Bright-field images were taken after aequorin imaging. The total remaining aequorin luminescence (L_{max}) was estimated by discharging with 0.9 M $CaCl_2$ in 10% (v/v) ethanol[9,17]. The calibration of $[Ca^{2+}]_i$ measurements was performed as described previously ($pCa = 0.6747 \times (-\log k) + 5.3177$, where k is a rate constant equal to luminescence counts (L) divided by total remaining counts (L_{max})[17]. All treatments were carried out in the dark, and the experiments were carried out at room temperature (22–24 °C).

Generation of T-DNA insertion-mutagenized *Arabidopsis* populations. The T-DNA insertion-mutagenized *Arabidopsis* populations for genetic screens[52] were generated using wild-type *Arabidopsis* Col-0 expressing aequorin[9]. Initially, we used the activation tagging pSKI015 vector[53] and the flower dipping method[54] to transform aequorin-expressing *Arabidopsis*, and generated T-DNA insertion populations of over 100,000 lines. Unfortunately, a very high percentage (>30%) of these T-DNA lines did not show a wild-type level of aequorin bioluminescence signals by discharging with 0.9 M $CaCl_2$, suggesting that the aequorin 35S promoter was largely silenced, as reported previously[55]. Therefore, these populations could not be used to screen for mutants defective in stimulus-induced increases in $[Ca^{2+}]_i$. Subsequently, we decided to use a different vector without the 35S promoter, the pBIB-BASTA vector (a gift from J. Li)[56]. We did not find a high percentage of aequorin silencing in these aequorin-expressing *Arabidopsis* lines transformed with pBIB-BASTA. Then, we regenerated the T-DNA insertion populations, and collected the seeds for over 120,000 independent T2 transgenic lines in pools.

Screens for mutants defective in salt-induced increases in $[Ca^{2+}]_i$ (*moca*). The aequorin Ca^{2+}-imaging-based genetic screens for mutants defective in monocation-induced $[Ca^{2+}]_i$ increases (*moca*) were carried out largely as described previously for the *reduced hyperosmolality-induced $[Ca^{2+}]_i$ increases* (*osca*) mutants[17]. In brief, T3 seeds were sterilized, and individual seeds were planted evenly using a template in 150 mm × 15-mm Petri dishes (260 seeds per Petri dish), and grown in ½ MS medium for seven days. Aequorin bioluminescence images were acquired for the salt treatment, that is, addition of 100 ml 200 mM NaCl solution into Petri dishes via a home-made device. A total of ~86,000 T3 seeds was screened in the first round, and ~10,000 seedlings that showed weaker $[Ca^{2+}]_i$ increases were picked up. These seedlings were then transferred to soil, and collected individually for seeds. From the second- to the fourth-round screens, individual lines were analysed for the *moca* phenotype and six putative mutants with a stable *moca* phenotype were isolated. To ensure that the *moca* phenotype was not caused by defects related to aequorin-based Ca^{2+} imaging, we sequenced the aequorin transgene in these putative mutants to eliminate those lines with T-DNA insertion or mutation in aequorin. We also analysed the total aequorin signals by discharging to eliminate those lines in which aequorin expression was silenced. In addition, we paid extra attention to those lines with growth and developmental phenotypes, such as small in size and dark leaf colour, and in general selected those with no obvious growth and developmental phenotypes for further studies.

Yellow Cameleon-based $[Ca^{2+}]_i$ imaging in root cells. The *moca1* mutant was crossed into wild-type plants constitutively expressing the Ca^{2+} sensor YC3.6, and GFP FRET-based Ca^{2+} imaging was carried out as described[17,51,57]. More than ten homozygous lines were generated and analysed. Cameleon-based measurements of $[Ca^{2+}]_i$ spikes and waves in root cells were conducted as described previously[21,22]. In brief, *Arabidopsis* seedlings expressing YC3.6 were grown under in a thin layer (~2 mm) of 1.0% (w/v) agar and ½ MS medium on a cover glass for five days under light. A small window (~1 mm × 1 mm) was removed from the agar gel at the tip of the root to expose the apical ~500 μm of the root and allow accurate application of NaCl solution. For inhibitor treatments, a small window (~1 mm × 1 mm) was made in the gel in the middle region of the root (Fig. 3g). Five to ten microlitres of 50 μM $LaCl_3$ was added in the middle gel window about 30 min before salt treatment. Ratiometric Ca^{2+} imaging was performed using a fluorescence microscope (Axiovert 200/Axio Observer 3; Zeiss) equipped with two filter wheels (Lambda 10-2/10-3; Sutter Instruments), and a cooled CCD camera or CMOS (CoolSNAP$_{fx}$/Prime sCMOS; Roper Scientific)[17]. Excitation was provided at 440 nm and emission ratiometric ($F_{535\,nm}/F_{485\,nm}$) images were collected using MetaFluor software.

Growth responses to salt stress. For the salt tolerance assay, wild-type and *moca1* mutant *Arabidopsis* seeds were sterilized, kept in darkness at 4 °C for three days, and then placed in low calcium (0.2 mM $CaCl_2$) ½ MS agar medium as described[58], containing different concentrations of NaCl. The ½ MS agar medium contained major salts (NH_4NO_3, $MgSO_4$, KH_2PO_4, and KNO_3; Sigma), vitamin solution (Sigma), 0.5% sucrose (Sigma), 0.6% agar (Sigma), 0.05% MES (Sigma), and calcium supplemented to 0.2 mM $CaCl_2$. Seedlings were grown for 12 days, and seedling weight, root length, and survival rate were analysed.

Na and K content analysis by ICP-MS. The content of cations was analysed using inductively coupled plasma mass spectrometry (ICP-MS) as described previously[59–61]. *Arabidopsis* seedlings were grown in Petri dishes under several NaCl concentrations for two weeks as described above, and harvested for the analysis of Na and K contents. Seedlings were dried overnight at 65 °C and weighed before digestion. Samples were digested in 75% nitric acid and 25% hydrogen peroxide for 3 h at 180 °C in a microwave digestion system (Ethos One, Milestone). Each sample was diluted to 10.0 ml with 18 MΩ water and measured using a NexIon300X ICP-MS Spectrometer (PerkinElmer, USA).

Na^+/H^+ exchange assay. The activities of Na^+/H^+ exchange were analysed as described previously[62]. *Arabidopsis* seedlings were grown in flasks in liquid medium (low Ca^{2+} and ½ MS) on a shaker for 12 d. One day before harvesting, the original medium was replaced with fresh medium supplemented with or without 100 mM NaCl. Plasma membrane-enriched vesicles were prepared using aqueous two-phase (Dextron-PEG3350) partitioning[62]. Inside-out vesicles were produced by adding 0.05% (w/v) Brij58 to the medium[63]. The protein content was determined by Bradford's method using BSA as a standard[64]. The membrane identity and transport competence of the vesicles were assessed by measuring the H^+-transport activity of the plasma membrane H^+-ATPase[62]. Na^+/H^+ exchange activity was measured as a Na^+-induced dissipation of ΔpH[62,65]. When the maximum ΔpH was formed (reached steady state), NaCl was added to initiate Na^+ transport. To determine initial rates of Na^+/H^+ exchange (change in fluorescence per minute, $\Delta\%F$/min), changes in relative fluorescence were measured 15 s after the addition of Na^+. Specific activity was calculated by dividing the initial rate by the mass of plasma membrane protein in the reaction ($\Delta\%F$/min per mg of protein)[62].

Genetic analysis and physical mapping. We could not identify the T-DNA insertion in the *moca1* mutant by either adaptor ligation-mediated PCR or thermal asymmetric interlaced PCR (Tail PCR). We also found that the *moca1* mutant had lost Basta resistance and could not survive the Basta selection in the same way as Col-0 (aequorin). We back-crossed *moca1* to aequorin-expressing Col-0 (aequorin) plants. The homozygous *moca1* lines in the Col-0 (aequorin) background, which showed a 1:3 mutant:wild-type ratio, were crossed to the ecotype Wassilewskija (Ws) followed by self-pollinating F_1 progeny to yield an F_2 population as described previously for physical mapping[17]. For *moca1* mapping, a total of ~5,200 F_2 seedlings (out of ~6,800 seeds) grown on Petri dishes that showed kanamycin resistance (aequorin transgene) were transferred to soil. We then genotyped aequorin using PCR, and aequorin homozygous lines were harvested individually for F_3 seeds. These F_3 lines were analysed individually for the *moca1* phenotype using aequorin imaging. Eventually, homozygous *moca1* lines with homozygous aequorin were obtained as the mapping population. Linkage analysis of F_2 plants revealed that the *moca1* locus was located on chromosome 5. Markers for fine mapping were searched from the databases of https://www.arabidopsis.org and http://archive.is/amp.genomics.org.cn/. These markers were used to perform PCR and to isolate the interval that flanks the mutation as described previously[17]. Finally, we fine-mapped the mutation into the narrowest interval, then sequenced open reading frames (ORFs) in the interval, and identified a 12-base-pair deletion in an ORF in *moca1*. Transmembrane α-helical spanners of MOCA1 were predicted by various models using Aramemnon (http://aramemnon.botanik.uni-koeln.de)[66].

DNA constructs and transgenic lines. Gateway cloning[67] was used to construct *p35S::MOCA1*, *p35S::MOCA1–GFP*, *pMOCA1::GUS*, *pMOCA1::MOCA1–GFP*, and *pMOCA1::mMOCA1–GFP*. The *MOCA1* full-length complementary or

genomic DNA and the 2-kb promoter region were amplified by PCR from cDNA and genomic DNA, respectively. The DNA fragment and the promoter region were cloned into the pENTR vector (Invitrogen). Coding sequences were transferred from the entry clones to gateway-compatible destination vectors. Transgenic *Arabidopsis* lines were generated by *Agrobacterium*-mediated transformation[54], and homozygous transgenic T3 lines carrying a single insertion were used. The *moca1-2* (SALK_131321) and *moca1-3* (GK-856G03) lines were obtained from the *Arabidopsis* Biological Resource Center (ABRC). Both insertions are located within the gene. *moca1-2* is found in the fourth exon, and *moca1-3* is found in the eighth exon. We were unable to obtain homozygotes for either mutant allele. Note that *moca1* is *moca1-1* in this study.

MOCA1–GFP subcellular localization analysis. For analysis of MOCA1–GFP in *Arabidopsis* seedlings, MOCA1 promoter-driven wild-type MOCA1 (*pMOCA1::MOCA1–GFP*; full-length genomic DNA) and mutant MOCA1 (*pMOCA1::mMOCA1–GFP*, full-length genomic DNA) and 35S promoter-driven MOCA1 (*p35S::MOCA1–GFP*, full-length complementary DNA) transgenic plants were generated as described previously[17]. Seedlings grown in ½ MS medium in Petri dishes for seven days were subjected to GFP confocal imaging with the Zeiss LSM 510 microscope or whole-seedling imaging with a Zeiss SteREO Discovery V20/V16 microscope. Data represent more than ten independent lines examined, which displayed similar GFP subcellular localization. We analysed the fluorescence in the *pMOCA1::MOCA1–GFP* and the Golgi maker *mCherry–Golgi* co-transgene lines[68], and observed that MOCA1–GFP displayed punctate patterns in the cytosol, and was co-localized with the mCherry–Golgi marker. The Golgi localization is consistent with the prediction by SUBA4 (http://suba.live/)[69], and well supported by several previous studies on Golgi proteomes[70,71].

Histochemical GUS activity analysis. Histochemical staining for GUS activity using the MOCA1 promoter-driven GUS (*pMOCA1::GUS*) transgenic lines was performed as described previously[17]. Seedlings grown in ½ MS medium or soil were used for histochemical staining[72]. Data represent six independent lines examined, which displayed similar staining patterns.

PCR primers and vectors. Genotyping primers: *MOCA1-LP*, 5′-CATTC TTGTCCTTAT AGTTGCTGGT; *MOCA1-RP*, 5′-CTTTCAAGAACCATCTCA CCGC. Cloning primers: *MOCA1-cDNA_Fw*, 5′-CACCATGGTG AGACTC AAGACGAGT; *MOCA1-cDNA_Rev*, 5′-(TCA)ACAGAGGAAACATA GGGAAT; Primers for *MOCA1* promoter: *MOCA1P_Fw*, 5′-CACCGGATAATGTT GAGTAATGG; *MOCA1P_Rev*, 5′-ACCTTTGTTCCTTGTCACCG. Vectors: *p35S::MOCA1:pGWB502Ω*, *pMOCA1::GUS:pGWB533*, *p35S::MOCA1-GFP:pGWB405*, *pMOCA1::MOCA1-GFP:pGWB405*, *pMOCA1::MOCA1-GFP:pGWB404*, and *pMOCA1::mMOCA1-GFP:pGWB404*.

Sphingolipid analysis. *Arabidopsis* GIPCs were extracted and purified from 7-day-old seedlings as described previously with modifications[23,50,73]. In brief, seedlings (~20 g fresh weight) were blended with 300 ml cold 0.1 N aqueous acetic acid. The slurry was filtered through eight layers of miracloth, and the residue was then re-extracted with hot (70 °C) 70% ethanol containing 0.1 N HCl. The filtrates were left immediately at −20 °C overnight. The precipitate was pelleted by centrifugation at 2,000*g* at 4 °C. The GIPC-containing pellet was washed with cold acetone, and subsequently with cold diethyl ether to yield a whitish precipitate. The GIPC crude extracts were dissolved in tetrahydrofuran (THF)/methanol/water (4:4:1, v/v/v) containing 0.1% formic acid at 60 °C, further dried, and submitted to a butan-1-ol/water (1:1, v/v) phase partition. The upper butanolic phase was dried and the residue was dissolved in THF/methanol/water (4:4:1, v/v/v) containing 0.1% formic acid. The GIPC solutions were analysed using a saturated solution of 2,6-dihydroxy-acetophenone (DHA) matrix that was prepared in 50:50 (v/v) ethanol/water containing 3 mM ammonium sulfate. The concentration of GIPC samples was ~0.5 mg/ml in THF/methanol/water (4:4:1, v/v/v) containing 0.1% formic acid. Samples were mixed with matrix solution (0.5 μl) before being loaded on the MALDI plate. Spectra were acquired in negative ion mode on a MALDI Q-TOF mass spectrometer (AB SCIEX TOF/TOF 5800). The laser was set to an energy level of 250 on the instrument scale. The mass range was from m/z 800 to 3,000. The diagram of plant GIPC structure shown in Extended Data Fig. 8h was adapted from previous studies[23,27,30].

Protoplast isolation and ζ potential measurement. The ζ potential measurements of *Arabidopsis* mesophyll protoplasts were carried out using similar methods to those described previously[74–76]. Upper epidermis samples from four-week-old *Arabidopsis* leaves were vacuum infiltrated for 10 min and then incubated for 90 min in enzyme solutions containing 1% (w/v) cellulase (Onozuka R-10), 0.4% (w/v) macerozyme (Onozuka R-10), 0.5 M mannitol, and 5 mM MES-KOH, pH 5.7. Released protoplasts were filtered through a 100-μm nylon mesh and washed twice in solution without enzymes[77,78]. For the ζ potential measurements, protoplasts in 2 ml solution containing several concentrations of NaCl were loaded into a one-well chambered cover glass (Nunc) with installed integral platinum electrodes at the end of the chamber. The migration of the protoplasts under a potential gradient of about 10 V/cm was analysed[74] using a fluorescence microscope (Axiovert 200/Axio Observer 3) equipped with filter wheels and a cooled CCD camera[17]. More than five measurements for each sample, each consisting of 30 runs with duration of ~30 s, were performed at room temperature. For each concentration of NaCl, at least three independent preparations were analysed. An average of the electrophoretic velocity value was analysed and calculated using Image J (https://imagej.nih.gov/ij/). The values of ζ potentials were further calculated using the Helmholtz–Smoluchowski equation[74–76]. For survival rate, protoplasts were treated with NaCl, and white light images were taken every 30 s for 20 min to monitor the integrity of protoplasts (Fig. 5f).

Isothermal titration calorimetry. ITC[79–82] was used to analyse binding of Na^+ to GIPCs with modifications[83,84]. The purified lipids were titrations in running buffer containing 20 mM MES-Tris (pH 5.8) and 240 mM mannitol. For analysis of binding of Na^+ to GIPCs, small unilamellar vesicles of GIPCs from the wild-type and *moca1* seedlings were produced by sonication in running buffer (30 min on ice, 20 s pulse on, 20 s pulse off, amplitude 25%)[50]. Lipid vesicles were produced from the MS-analysed wild-type GIPC–IPC mixture (calculated as 5.792:0.608 wt:wt) or *moca1* GIPC–IPC mixture (calculated as 0.416:5.984 wt:wt) at a final lipid concentration of 6.40 mg/ml, and the concentrations of the wild-type and *moca1* GIPC–IPC mixtures were 5.21 mM and 6.77 mM, respectively. ITC measurements were performed at 25 °C using MicroCal iTC200 (Malvern Panalytical)[79–82]. NaCl at a concentration of 50 mM was injected into a 200 μl sample cell until saturation was reached. The volume of each injection was 2 μl with a total of 19 injections and consecutive injections were separated by 2 min to allow the peak to return to baseline. Similarly, binding of K^+ or Li^+ to GIPCs was analysed with the experimental procedures developed for Na^+. ITC data were analysed with a one-site fitting model using Origin software (Malvern Panalytical and OriginLab). Error was calculated from the standard deviation of six titrations.

Phylogenetic analysis. Multiple sequence alignment of the GT8 and UDPGP regions was performed using MAFFT v7.05 with automatic method[85]. The phylogeny was constructed using FastTree v2.1.7 with default parameters[86]. FastTree implements an ultrafast and fairly accurate approximate ML method. Phylogenetic trees were represented and edited using FigTree (http://tree.bio.ed.ac.uk/software/figtree/). Statistical values are shown beside selected major nodes with black circles. The scale bar indicates the number of amino acid residue substitutions per site. The full tree could be divided into three major classes (I to III), consistent with previous results[87].

Statistical analysis. To minimize the system variations, wild-type and *moca1* seedlings as well as transgenic lines were always grown side-by-side in agar medium in Petri dishes or in soil in trays, and the Petri dishes and trays were rotated every other day in positions in the growth chambers or rooms to have even temperature and light. Individual seedlings or pools of several seedlings were analysed. For instance, in Fig. 2d, e, 8–12 seedlings were pooled together as one pool, and 12 of these pools were analysed to give mean ± s.d. Independent experiments were performed at least three times. Statistical analysis was performed using Excel 2016 software (Microsoft), and P values were calculated via T.TEST (Student's t-test, two-sided). Data are presented as mean ± s.d. or s.e.m. To analyse the difference between genotypes or treatments in line graphs, two-way analysis of variance (ANOVA) was carried out using SAS 9.3/9.4 software (SAS Institute). Values of $P < 0.05$ were considered statistically significant.

No statistical methods were used to predetermine sample size. The experiments were not randomized and investigators were not blinded to allocation during experiments and outcome assessment.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The data that support the findings of this study are available from the corresponding author upon request.

51. Monshausen, G. B., Messerli, M. A. & Gilroy, S. Imaging of the Yellow Cameleon 3.6 indicator reveals that elevations in cytosolic Ca^{2+} follow oscillating increases in growth in root hairs of *Arabidopsis*. *Plant Physiol*. **147**, 1690–1698 (2008).
52. Feldmann, K. A. T-DNA insertion mutagenesis in *Arabidopsis* - mutational spectrum. *Plant J*. **1**, 71–82 (1991).
53. Weigel, D. et al. Activation tagging in *Arabidopsis*. *Plant Physiol*. **122**, 1003–1013 (2000).
54. Clough, S. J. & Bent, A. F. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J*. **16**, 735–743 (1998).
55. Schubert, D. et al. Silencing in *Arabidopsis* T-DNA transformants: the predominant role of a gene-specific RNA sensing mechanism versus position effects. *Plant Cell* **16**, 2561–2572 (2004).
56. Li, J. et al. BAK1, an *Arabidopsis* LRR receptor-like protein kinase, interacts with BRI1 and modulates brassinosteroid signaling. *Cell* **110**, 213–222 (2002).
57. Pei, Z.-M. et al. Calcium channels activated by hydrogen peroxide mediate abscisic acid signalling in guard cells. *Nature* **406**, 731–734 (2000).

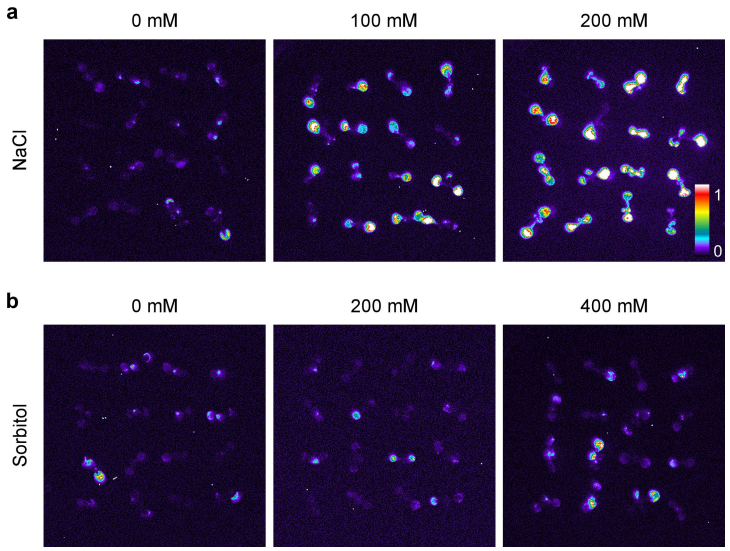
58. Tang, R. H. et al. Coupling diurnal cytosolic Ca^{2+} oscillations to the CAS-IP_3 pathway in *Arabidopsis*. *Science* **315**, 1423–1426 (2007).
59. Lahner, B. et al. Genomic scale profiling of nutrient and trace elements in *Arabidopsis thaliana*. *Nat. Biotechnol*. **21**, 1215–1221 (2003).
60. Baxter, I. et al. A coastal cline in sodium accumulation in *Arabidopsis thaliana* is driven by natural variation of the sodium transporter AtHKT1;1. *PLoS Genet*. **6**, e1001193 (2010).
61. Song, W. Y. et al. Arsenic tolerance in *Arabidopsis* is mediated by two ABCC-type phytochelatin transporters. *Proc. Natl Acad. Sci. USA* **107**, 21187–21192 (2010).
62. Qiu, Q. S., Guo, Y., Dietrich, M. A., Schumaker, K. S. & Zhu, J. K. Regulation of SOS1, a plasma membrane Na^+/H^+ exchanger in *Arabidopsis thaliana*, by SOS2 and SOS3. *Proc. Natl Acad. Sci. USA* **99**, 8436–8441 (2002).
63. Johansson, F., Olbe, M., Sommarin, M. & Larsson, C. Brij 58, a polyoxyethylene acyl ether, creates membrane vesicles of uniform sidedness. A new tool to obtain inside-out (cytoplasmic side-out) plasma membrane vesicles. *Plant J*. **7**, 165–173 (1995).
64. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem*. **72**, 248–254 (1976).
65. Blumwald, E. & Poole, R. J. Na/H antiport in isolated tonoplast vesicles from storage tissue of *Beta vulgaris*. *Plant Physiol*. **78**, 163–167 (1985).
66. Schwacke, R. et al. ARAMEMNON, a novel database for *Arabidopsis* integral membrane proteins. *Plant Physiol*. **131**, 16–26 (2003).
67. Karimi, M., Inzé, D. & Depicker, A. GATEWAY vectors for *Agrobacterium*-mediated plant transformation. *Trends Plant Sci*. **7**, 193–195 (2002).
68. Nelson, B. K., Cai, X. & Nebenführ, A. A multicolored set of *in vivo* organelle markers for co-localization studies in *Arabidopsis* and other plants. *Plant J*. **51**, 1126–1136 (2007).
69. Hooper, C. M., Castleden, I. R., Tanz, S. K., Aryamanesh, N. & Millar, A. H. SUBA4: the interactive data analysis centre for *Arabidopsis* subcellular protein locations. *Nucleic Acids Res*. **45**, D1064–D1074 (2017).
70. Heard, W., Sklenář, J., Tomé, D. F. A., Robatzek, S. & Jones, A. M. E. Identification of regulatory and cargo proteins of endosomal and secretory pathways in *Arabidopsis thaliana* by proteomic dissection. *Mol. Cell. Proteomics* **14**, 1796–1813 (2015).
71. Nikolovski, N. et al. Putative glycosyltransferases and other plant Golgi apparatus proteins are revealed by LOPIT proteomics. *Plant Physiol*. **160**, 1037–1051 (2012).
72. Jefferson, R. A., Kavanagh, T. A. & Bevan, M. W. GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J*. **6**, 3901–3907 (1987).
73. Buré, C. et al. Fast screening of highly glycosylated plant sphingolipids by tandem mass spectrometry. *Rapid Commun. Mass Spectrom*. **25**, 3131–3145 (2011).
74. Wang, P. et al. Cell membrane surface potential (ψ_0) plays a dominant role in the phytotoxicity of copper and arsenate. *Plant Physiol*. **148**, 2134–2143 (2008).
75. Delgado, A. V., González-Caballero, F., Hunter, R. J., Koopal, L. K. & Lyklema, J. Measurement and interpretation of electrokinetic phenomena. *J. Colloid Interface Sci*. **309**, 194–224 (2007).
76. Kinraide, T. B. & Wang, P. The surface charge density of plant cell membranes (σ): an attempt to resolve conflicting values for intrinsic σ. *J. Exp. Bot*. **61**, 2507–2518 (2010).
77. Yoo, S. D., Cho, Y. H. & Sheen, J. *Arabidopsis* mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nat. Protocols* **2**, 1565–1572 (2007).
78. Endo, M., Shimizu, H. & Araki, T. Rapid and simple isolation of vascular, epidermal and mesophyll cells from plant leaf tissue. *Nat. Protocols* **11**, 1388–1395 (2016).
79. Welsch, M. E. et al. Multivalent small-molecule pan-RAS inhibitors. *Cell* **168**, 878–889 (2017).
80. Menting, J. G. et al. How insulin engages its primary binding site on the insulin receptor. *Nature* **493**, 241–245 (2013).
81. Coscia, F. et al. Multi-level proteomics identifies CT45 as a chemosensitivity mediator and immunotherapy target in ovarian cancer. *Cell* **175**, 159–170 (2018).
82. Chen, Y. N. P. et al. Allosteric inhibition of SHP2 phosphatase inhibits cancers driven by receptor tyrosine kinases. *Nature* **535**, 148–152 (2016).
83. Johnson, R. A., Manley, O. M., Spuches, A. M. & Grossoehme, N. E. Dissecting ITC data of metal ions binding to ligands and proteins. *Biochim. Biophys. Acta*. **1860**, 892–901 (2016).
84. Machaidze, G., Ziegler, A. & Seelig, J. Specific binding of Ro 09-0198 (cinnamycin) to phosphatidylethanolamine: a thermodynamic analysis. *Biochemistry* **41**, 1965–1971 (2002).
85. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol*. **30**, 772–780 (2013).
86. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol*. **26**, 1641–1650 (2009).
87. Yin, Y., Chen, H., Hahn, M. G., Mohnen, D. & Xu, Y. Evolution and function of the plant cell wall synthesis-related glycosyltransferase family 8. *Plant Physiol*. **153**, 1729–1746 (2010).

Acknowledgements We thank J. Li for the pBIB-BASTA vector; M. R. Knight for aequorin-expressing seeds; S. Gilroy for YC3.6-expressing seeds; M. Chen, X. Dong and T.-p. Sun for advice on physical mapping; and T.-p. Sun, J. Pei, D. R. McClay and P. N. Benfey for discussion and reading of the manuscript. Z.J., M.T., F.W., W.Z. and H. Zhang were supported in part by the Shenzhen Peacock Innovative Research Team Program; Z.J., X.Z., L.L., X.W., F.W. and F.Y. by the Pandeng Project of Hangzhou Normal University and Chinese NSF (31301170, U1130304); and F.W., Y.N. and Y.X. by fellowships from the China Scholarship Council. This work was supported by grants from NSF (IOS-1457257, IOS-0848263) and USDA (CSREES-2006-35100-17304) to Z.-M.P.; National Key Project for Synthetic Biology (2018YFA0902500) to Z.H.; and DOE (DE-SC0014077) to T.V.-D. and Z.-M.P.

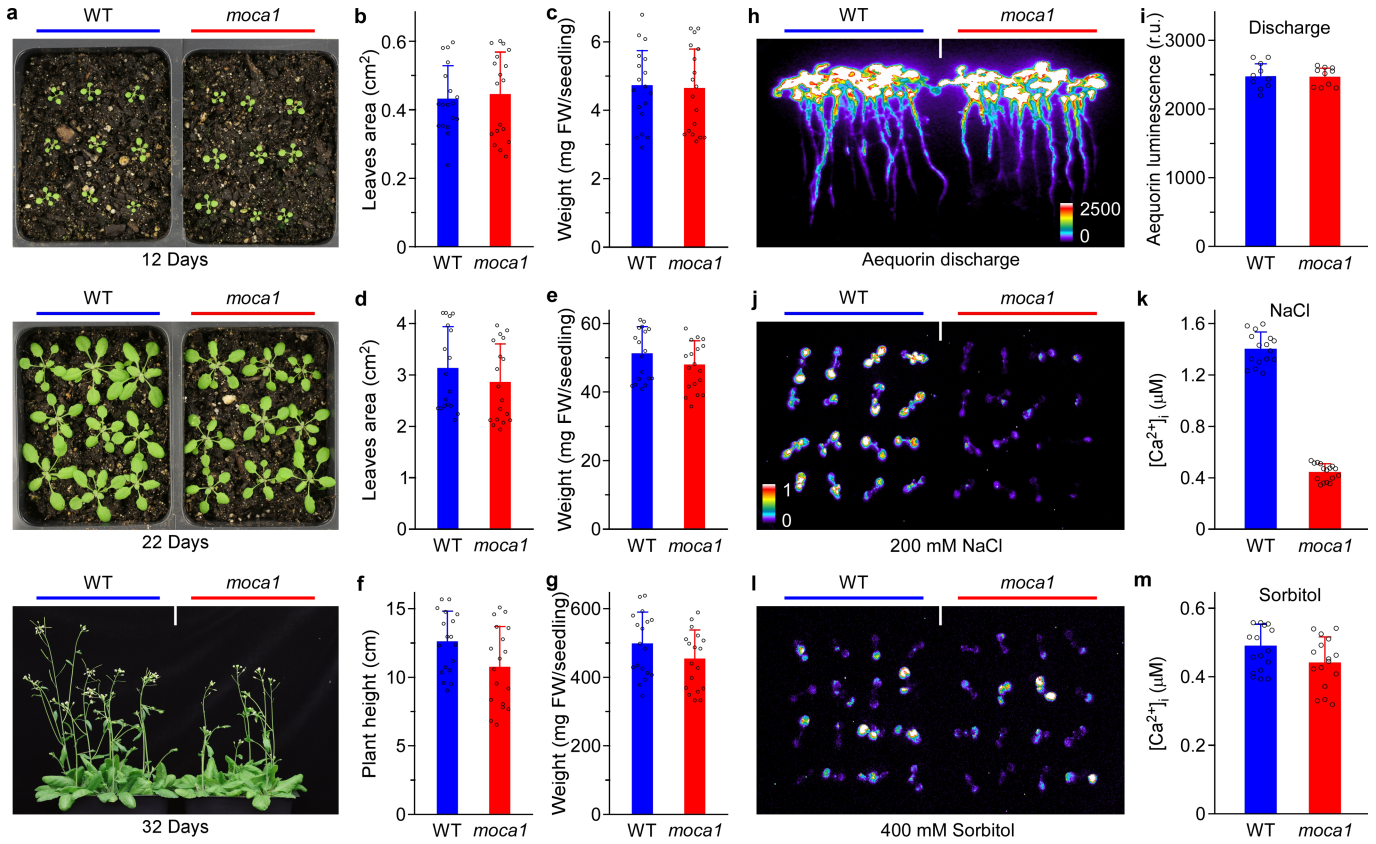
Author contributions Z.J., F.Y., Y. Xiang and Y.N. conducted genetic screens. Z.J., X.Z., F.Y., L.L., F.L., Y. Xue, F.W., X.W. and H. Zhao performed map-based cloning. C.L., R.Y., F.Y. and B.B. generated T-DNA-mutagenized *Arabidopsis* populations. Z.J., L.L., F.W., M.T., C.H., H. Zhao, C.P. and W.Z. carried out physiological analyses. Z.J., M.T., W.Z., and Z.H. analysed GIPC content and ion content. M.T., Z.J., W.Z., and Z.H. analysed ion binding to GIPCs. Z.J., C.P. and F.W. conducted zeta potential measurement. Z.J., L.L., H.-N.W., B.M.C., D.M.J., G.B.S., H. Zhang and T.V.-D. participated in Ca^{2+}-imaging analyses. Z.-M.P. designed and supervised the overall research with input from Z.H., Z.J. and J.N.S. Z.-M.P., Z.J. and J.N.S. wrote the paper. All authors discussed the results and commented on the manuscript.

Competing interests The authors declare no competing interests.

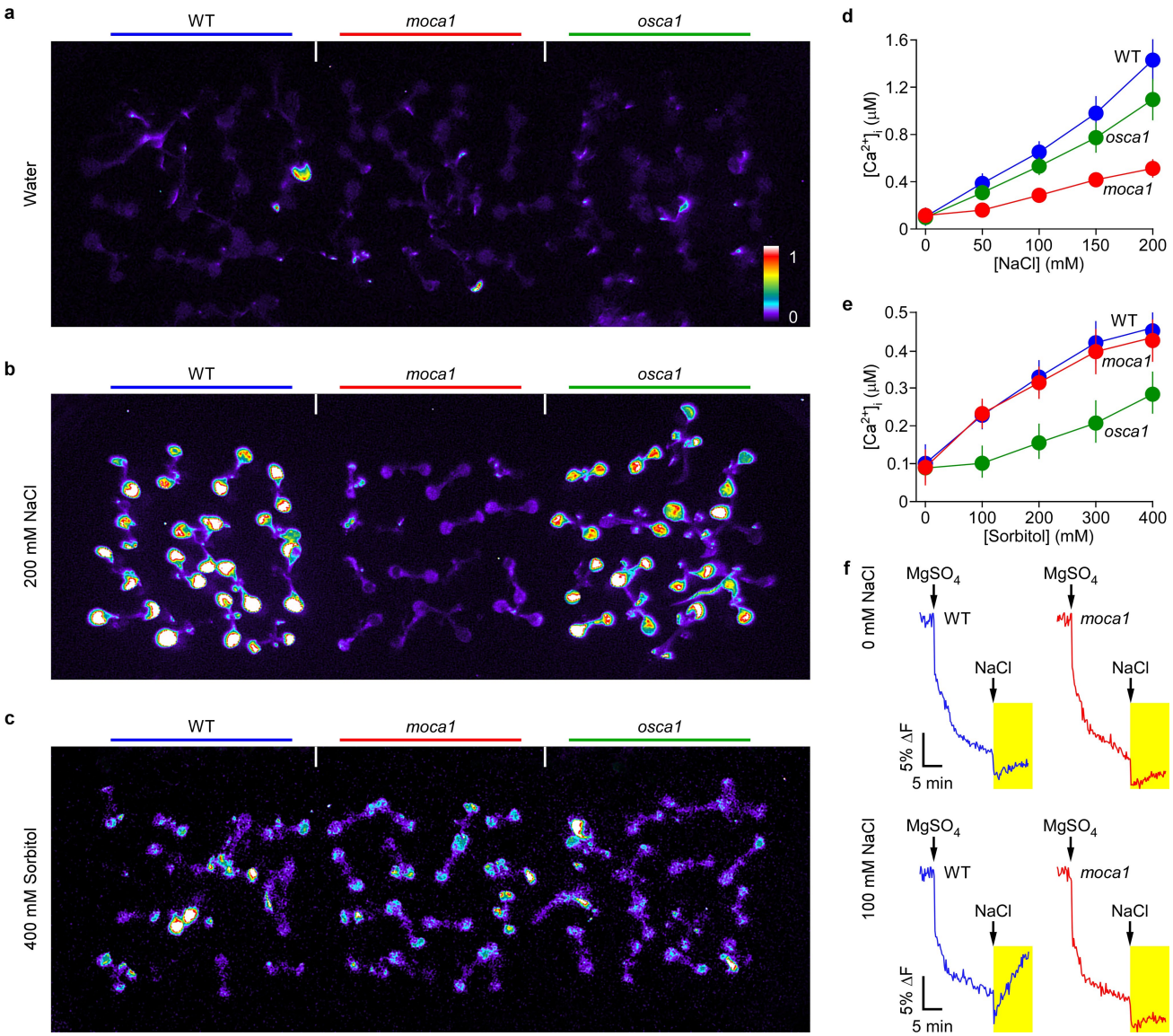
Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41586-019-1449-z.
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1449-z.
Correspondence and requests for materials should be addressed to Z.H. or Z.-M.P.
Reprints and permissions information is available at http://www.nature.com/reprints.



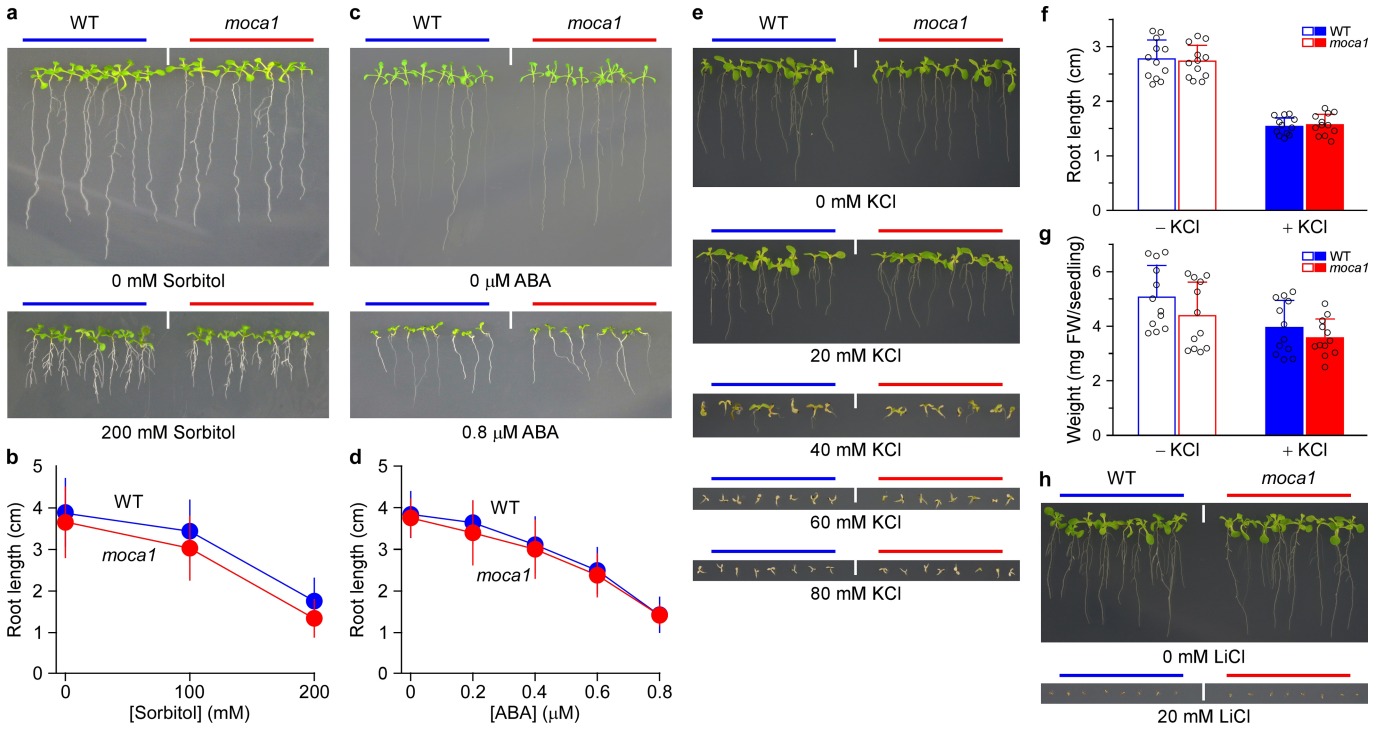
Extended Data Fig. 1 | The optimized conditions for genetic screens that distinguish increases in $[Ca^{2+}]_i$ induced by salt from those induced by osmotic stress. a, Aequorin bioluminescence imaging of $[Ca^{2+}]_i$ induced by NaCl at concentrations of 0, 100, or 200 mM in aequorin-expressing *Arabidopsis* seedlings. Relative $[Ca^{2+}]_i$ in leaves is shown using a pseudo-colour scale. Similar results were seen in more than 20 independent experiments. **b**, Aequorin imaging of $[Ca^{2+}]_i$ induced by sorbitol at concentrations of 0, 200, or 400 mM (at equivalent osmolalities to 0, 100, and 200 mM NaCl in **a**, respectively). Seedlings grown for 7 days were treated with NaCl or sorbitol solutions at indicated concentrations, and the aequorin bioluminescence images were acquired as described for experiments in Fig. 1a. Similar results were seen in more than 20 independent experiments.



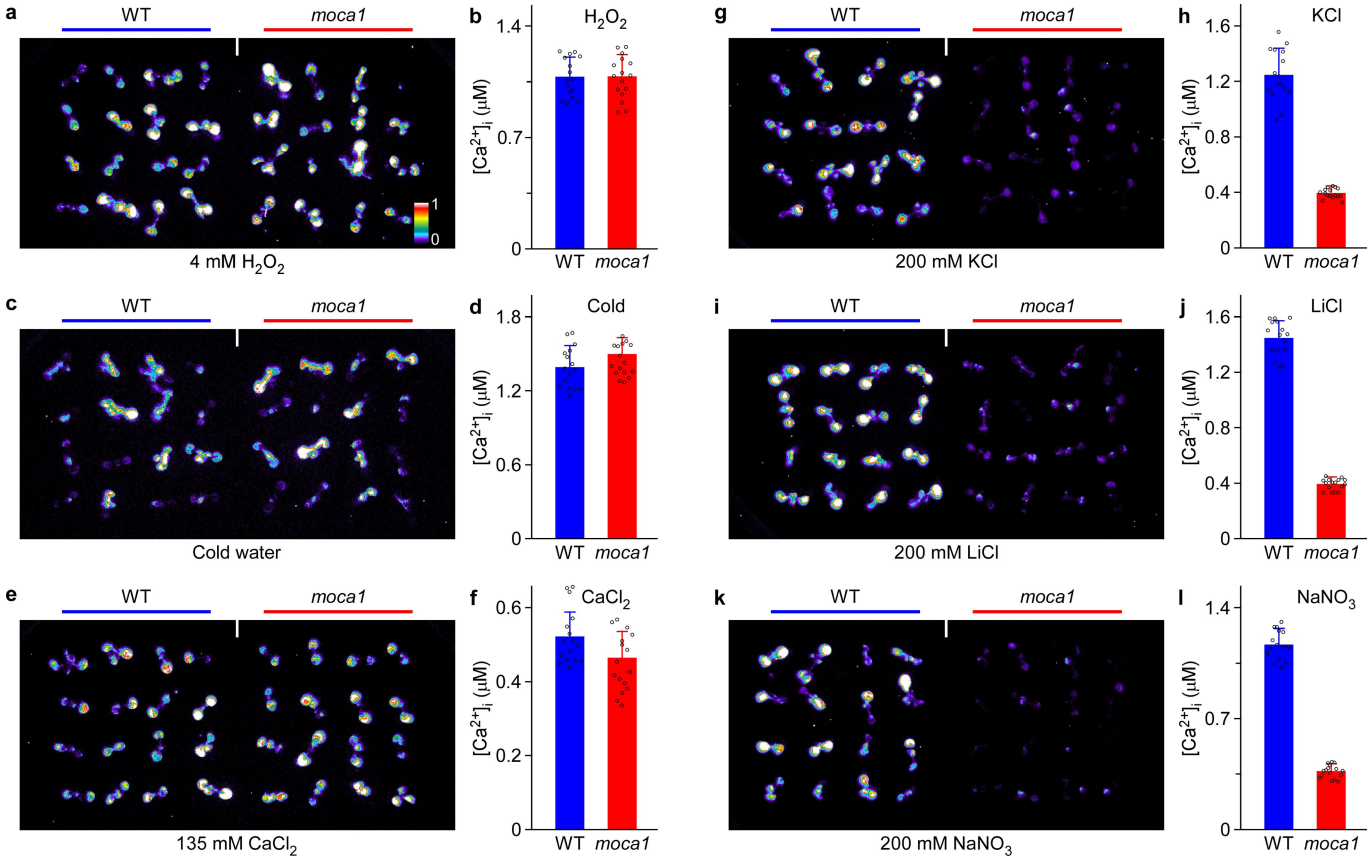
Extended Data Fig. 2 | The *moca1* mutant has wild-type-like growth phenotypes and shows defects in increases in $[Ca^{2+}]_i$ induced by salt but not by osmotic stress. **a**, Representative photographs of wild-type and *moca1* plants grown in soil taken at 12 days (top; young seedling stage), 22 days (middle; bolting stage), and 32 days (bottom; flowering and seed setting stage). Similar results from several batches of seeds were seen in more than 50 experiments. **b**, **c**, Leaf area (**b**) and shoot weight (**c**) from 12-day-old seedlings as in **a** (**b**, $P = 0.729$; **c**, $P = 0.912$). **d**, **e**, Leaf area (**d**) and shoot weight (**e**) from 22-day-old plants as in **a** (**c**, $P = 0.141$; **d**, $P = 0.179$). **f**, **g**, Plant height (**f**) and shoot weight (**g**) from 32-day-old plants as in **a** (**f**, $P = 0.054$; **g**, $P = 0.168$). **b**–**f**, Mean ± s.d.; $n = 18$ seedlings. **h**, **i**, Similar total amount of remaining aequorin bioluminescence in wild-type and *moca1* seedlings. The same seedlings used in Fig. 1b were treated with a solution containing 0.9 M $CaCl_2$ and 10% (v/v) ethanol to measure the total amount of remaining aequorin, and no difference between wild-type and *moca1* seedlings was observed (**h**). Similar results were seen in more than 30 separate experiments. Quantification of total amount of aequorin in wild-type and *moca1* plants from experiments as in **h** (**i**; mean ± s.d.; $n = 10$ pools (20 seedlings per pool); $P = 0.893$). **j**, **k**, The *moca1* mutant shows significantly lower elevation in $[Ca^{2+}]_i$ induced by 200 mM NaCl than the wild type. Wild-type and *moca1* seedlings grown side-by-side were treated with 200 mM NaCl solution (~400 mOsm) and changes in $[Ca^{2+}]_i$ were recorded (**j**). Data are quantified from experiments in **j** (**k**; mean ± s.d.; $n = 16$ seedlings; $P < 0.001$). Similar results were seen in more than 20 independent experiments. **l**, **m**, Similar increases in $[Ca^{2+}]_i$ induced by 400 mM sorbitol (with about equal osmolality to 200 mM NaCl) in wild-type and *moca1* seedlings. Wild-type and *moca1* seedlings grown side-by-side were treated with 400 mM (~400 mOsm) sorbitol solution and changes in $[Ca^{2+}]_i$ in leaves were recorded (**l**). Data are quantified from experiments in **l** (m; mean ± s.d.; $n = 16$ seedlings; $P = 0.141$). Similar results were seen in more than 20 independent experiments.



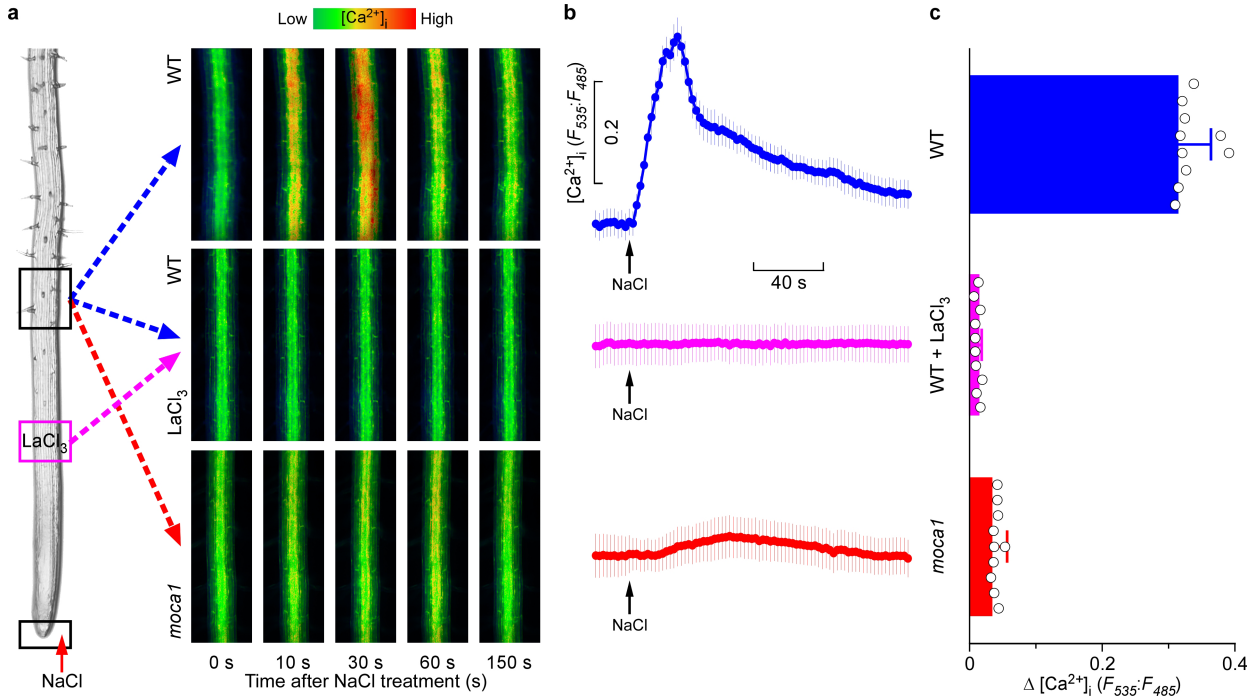
Extended Data Fig. 3 | Differences in Ca^{2+} signalling between salt stress-sensing *moca1* and osmotic stress-sensing *osca1* mutants and impaired activation of SOS1 Na^+/H^+ antiporter by salt in *moca1*. **a**–**c**, Representative aequorin bioluminescence images of wild-type, *moca1* and *osca1* seedlings grown side-by-side and treated with water (**a**), 200 mM NaCl (**b**), or 400 mM sorbitol (**c**). Similar results were seen in more than 20 independent experiments. **d**, **e**, Averaged increases in $[Ca^{2+}]_i$ plotted as a function of applied NaCl concentrations (**d**) and sorbitol concentrations (**e**) in wild-type, *moca1* and *osca1* leaves. Similar results were seen in more than ten separate experiments. Data are from three experiments (mean ± s.d.; $n = 48$ seedlings for each data point). **f**, Representative fluorescence traces of the pH-sensitive probe quinacrine show plasma membrane Na^+/H^+ exchange activity from wild-type and *moca1* plants without (0 mM NaCl) or with salt treatment (100 mM NaCl) for 24 h. Plasma membrane vesicles were prepared, and ΔpH was established by activation of the plasma membrane H^+-ATPase ($MgSO_4$) and measured as a decrease (quench) in the fluorescence of quinacrine (see Methods). Na^+/H^+ exchanger activity was measured as an increase in fluorescence (NaCl; highlighted). Quantification of the results is shown in Fig. 2f. Similar results were seen in three independent experiments.



Extended Data Fig. 4 | The *moca1* mutant does not affect responses to osmotic, KCl or LiCl stresses or the abscisic acid signalling pathway. **a**, **b**, Wild-type and *moca1* seedlings were grown in low Ca^{2+} and ½ MS medium containing 0 mM or 200 mM sorbitol for 12 days (**a**), and root lengths were analysed (**b**). Data are from three independent experiments (mean ± s.d.; $n = 13$ pools (8 seedlings per pool); two-way ANOVA, $P = 0.064$). **c**, **d**, Wild-type and *moca1* seedlings were grown in low Ca^{2+} and ½ MS medium in the presence or absence of abscisic acid (ABA) for 12 days (**c**), and root length was quantified (**d**). Data are from three independent experiments (mean ± s.d.; $n = 13$ pools (8–12 seedlings per pool); two-way ANOVA, $P = 0.249$). **e**–**g**, Wild-type and *moca1* seedlings were grown in low Ca^{2+} and ½ MS medium containing 0–80 mM KCl for 12 days (**e**), and root lengths (**f**) and fresh weight (**g**) from experiments in **e** were quantified. Data are from three independent experiments (mean ± s.d.; $n = 12$ pools (8–12 seedlings per pool); **f**, $P = 0.773$ (–KCl), $P = 0.707$ (+KCl); **g**, $P = 0.126$ (–KCl), $P = 0.281$ (+KCl)). **h**, Wild-type and *moca1* seedlings were grown in low Ca^{2+} and ½ MS medium containing 0 mM or 20 mM LiCl for 12 days. Severe inhibition of growth was observed in both wild-type and *moca1* seedlings at 20 mM LiCl. At low concentrations of LiCl, no differences were observed. Similar results were seen in more than 10 independent experiments. In contrast to NaCl stress (Fig. 2c–e), KCl and LiCl stresses inhibited the growth of wild-type and *moca1* seedlings at almost the same level (Extended Data Fig. 4e–h). Although the initial increases in $[Ca^{2+}]_i$ and activation of the SOS pathway are almost identical for short-term Na^+, K^+, and Li^+ stresses (Figs. 1e, 3d–f), the higher selectivity of SOS1 for Na^+ over K^+ and Li^+ might lead to higher Na^+ buildup but similar K^+ and Li^+ buildup in *moca1* seedlings compared to wild-type seedlings. Note also that, in natural environments, plants are exposed to NaCl stress often but KCl and LiCl stress rarely, implying that the hypersensitivity of *moca1* seedlings to NaCl stress is physiologically relevant.



Extended Data Fig. 5 | Selectivity of increases in $[Ca^{2+}]_i$ in response to various abiotic stimuli in the *moca1* mutant. **a**, **b**, Aequorin bioluminescence image of seedlings treated with 4 mM H_2O_2 (**a**) and quantification of $[Ca^{2+}]_i$ in leaves from experiments in **a** (**b**; $P = 0.982$). Aequorin images were acquired similarly to experiments in Fig. 3a. **c**, **d**, Aequorin bioluminescence image of seedlings treated with 12 °C cold water (**c**) and quantification of $[Ca^{2+}]_i$ in leaves from experiments in **c** (**d**; $P = 0.286$). Aequorin images were acquired similarly to experiments in Fig. 3b. **e**, **f**, Aequorin bioluminescence image of seedlings treated with 135 mM $CaCl_2$ (~400 mOsm) (**e**) and quantification of $[Ca^{2+}]_i$ in leaves from experiments in **e** (**f**; $P = 0.057$). Aequorin images were acquired similarly to experiments in Fig. 3c. **g**, **h**, Aequorin bioluminescence image of seedlings treated with 200 mM KCl (~400 mOsm) (**g**) and quantification of $[Ca^{2+}]_i$ in leaves from experiments in **g** (**h**; $P < 0.001$). Aequorin images were acquired similarly to experiments in Fig. 3d. **i**, **j**, Aequorin bioluminescence image of seedlings treated with 200 mM LiCl (~400 mOsm) (**i**) and quantification of $[Ca^{2+}]_i$ in leaves from experiments in **i** (**j**; $P < 0.001$). Aequorin images were acquired similarly to experiments in Fig. 3e. **k**, **l**, Aequorin bioluminescence image of seedlings treated with 200 mM $NaNO_3$ (~400 mOsm) (**k**) and quantification of $[Ca^{2+}]_i$ in leaves from experiments in **k** (**l**; $P < 0.001$). Aequorin images were acquired similarly to experiments in Fig. 3f.
All data show mean ± s.d.; $n = 16$ seedlings.



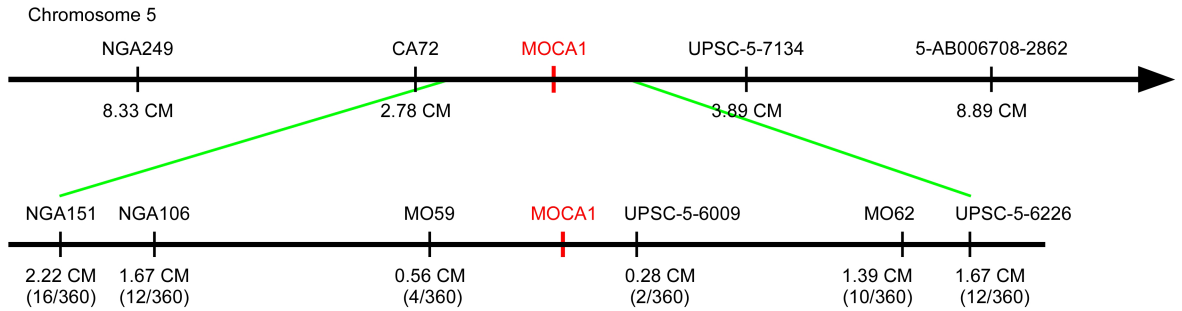
Extended Data Fig. 6 | Effect of Ca^{2+} channel blockers on salt-induced wave-like propagation of $[Ca^{2+}]_i$ in wild-type and *moca1* roots. a, Wave-like propagation of $[Ca^{2+}]_i$ increases through the root. Increases in $[Ca^{2+}]_i$ were monitored in ROIs 1,000 μm from the site of local application of 200 mM NaCl to root tip. Top, wild-type; middle, wild-type with 10 min pretreatment with 50 μM $LaCl_3$; bottom, *moca1* roots. YC3.6 Ca^{2+} imaging was carried out similarly to experiments in Fig. 3g–i. Similar results were seen in more than 20 independent experiments. **b**, Quantitative analyses of time course of $[Ca^{2+}]_i$ changes in ROIs in response to local treatment with 200 mM NaCl in similar experiments as in **a** (mean ± s.d.; $n = 10$ ROIs). **c**, Peak ratio changes from experiments similar to **a**, **b** (mean ± s.d.; $n = 10$ ROIs).

a

Crosses	Number of F2 plants		Ratio (WT : *moca1*)
	WT	*moca1*	
moca1 x Col_{AQ}[a]	203.3 ± 5.3	67.7 ± 3.5	3.0 ± 0.1
moca1 x Ws[b]	97.3 ± 3.8	33.2 ± 2.8	2.9 ± 0.2

Note: Col_{AQ}, *Arabidopsis* ecotype Col-0 expressing aequorin; Ws, *Arabidopsis* ecotype Wassilewskija; [a], *n* = 6 crosses; [b], *n* = 6 crosses.

b



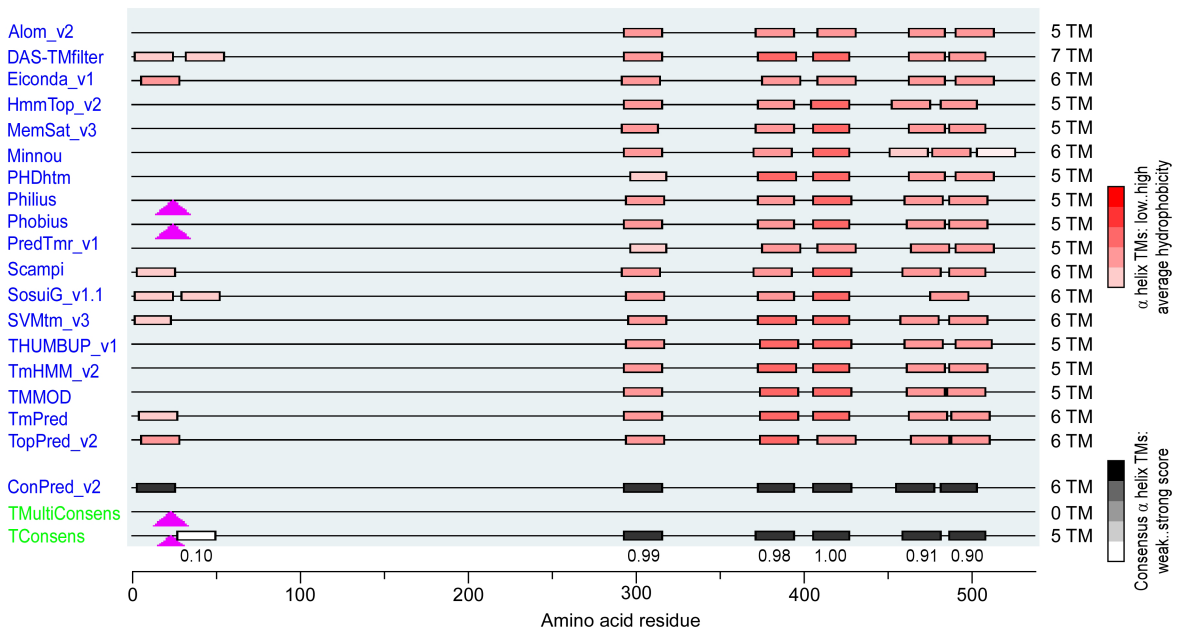
c

Marker name	Physical position (bp)	Primers (forward, reverse)	Type
MO59	5906000	AATAATGCAAAGTGCAAACGG, CGTTGAAATTATAACCGCGAG	SSLP
MO62	6224500	TTGTTTACAAAACCCACTCCG, GGCTTGATATGAGATGCCAAA	SSLP

d

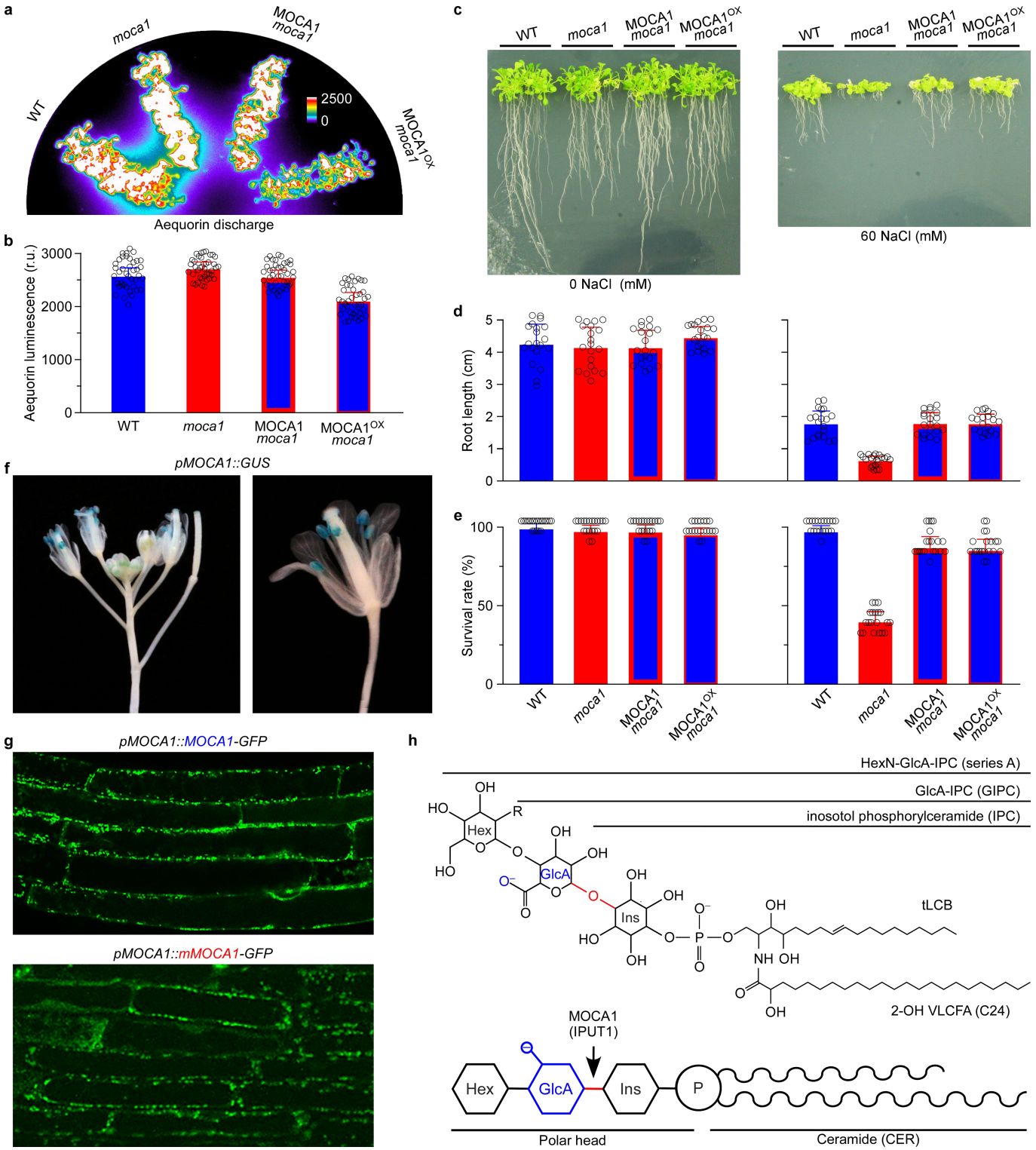
```
  1 MVRLKTSLWVLLLALVSIQLNGSFGSESSKVAYVTLLYGDEFLLGVRVLGKSIRDTGSTKDMVALVSDGVSDYSKKLLKADGWKVEKISLLANPNQVHPT
                TM1
101 RFWGVYTKLKIFNMTDYKKVVYLDADTIVVKNIEDLFKCSKFCANLKHSERLNSGVMVVEPSEALFNDMMRKVKTLSSYTGGDQGFLNSYYPDFPNARVF
201 DPSVTPEVLKTRPVPAMERLSTLYNADVGLYMLANKWMVDDSKLHVIHYTLGPLKPWDWWTAWLVKPVDAWHSIRVKLEETLPGTGGGSNQHDELVVKFL
301 FLLPLCALLFCIYRSIQGREGSLCWSSFSNQIRYLYYKVRSNGTLGYGGVSTMSPSYQPHSGNAQSKVPQHLGAVSVVVCFTAVLLSLGISFAIVPRQIM
      TM2                                                                    TM3      Δ 493-496
401 PWTGLVLVYEWTFTIFFLLFGVFLLFVHQHGKRIAIQSESSSLDDSAKVHQRAGGSCDVTTLYYGLGMAFLAIAAVSLPYILGITALFTRLGLMVGLAII
          TM4                                                    TM5                          TM6
501 LAAFMTYASEHLAVRWFLKGLEDRRDTTRSNSLCFLC
```

e



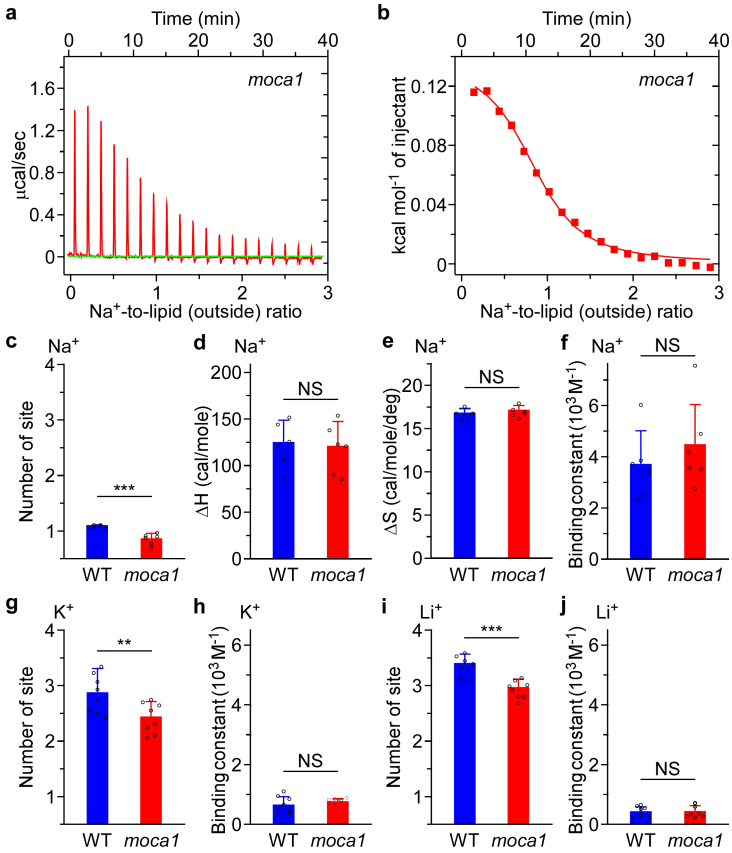
Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Genetic analysis and map-based cloning of *MOCA1*. **a**, All F_1 seedlings derived from *moca1* × wild-type (Col_{AQ}, Col-0 expressing aequorin) crosses showed wild-type salt-induced increases in $[Ca^{2+}]_i$. F_2 seedlings showed a 3:1 wild-type:*moca1* segregation, suggesting that the *moca1* phenotype resulted from a recessive mutation in a single nuclear gene. The F_2 seedlings, which were derived from *moca1* × Wassilewskija (Ws) crosses and also identified as aequorin homozygous, showed a 3:1 wild-type:*moca1* segregation. The same number of F_2 seeds for each cross were placed in Petri dishes and the phenotypes of salt-induced increases in $[Ca^{2+}]_i$ were scored for individual seedlings (mean ± s.e.m.; $n = 4$ for *moca1* × Col_{AQ} and *moca1* × Ws crosses). **b**, Physical mapping of *MOCA1*. *MOCA1* was positioned between NGA249 and 5-AB006708-2862 markers in the short arm of chromosome 5 in a segregating F_2 population derived from the *moca1* × Ws cross. *MOCA1* was fine-mapped to a region between MO59 and UPSC-5-6009 by analysing 720 recombinant chromosomes (360 lines) in the F_2 population with molecular markers described in **c**. We sequenced all open reading frames (ORFs) in this region between these two markers and identified one deletion in an ORF, which corresponded to the gene *At5g18480*. **c**, Molecular markers developed for fine mapping. **d**, *MOCA1* encodes a protein with six transmembrane α-helices (blue). Four amino acid residues from 493 to 496 (LMVG; red) are deleted in *moca1*. **e**, Transmembrane α-helical spanners predicted by various models using Aramemnon (http://aramemnon.botanik.uni-koeln.de).

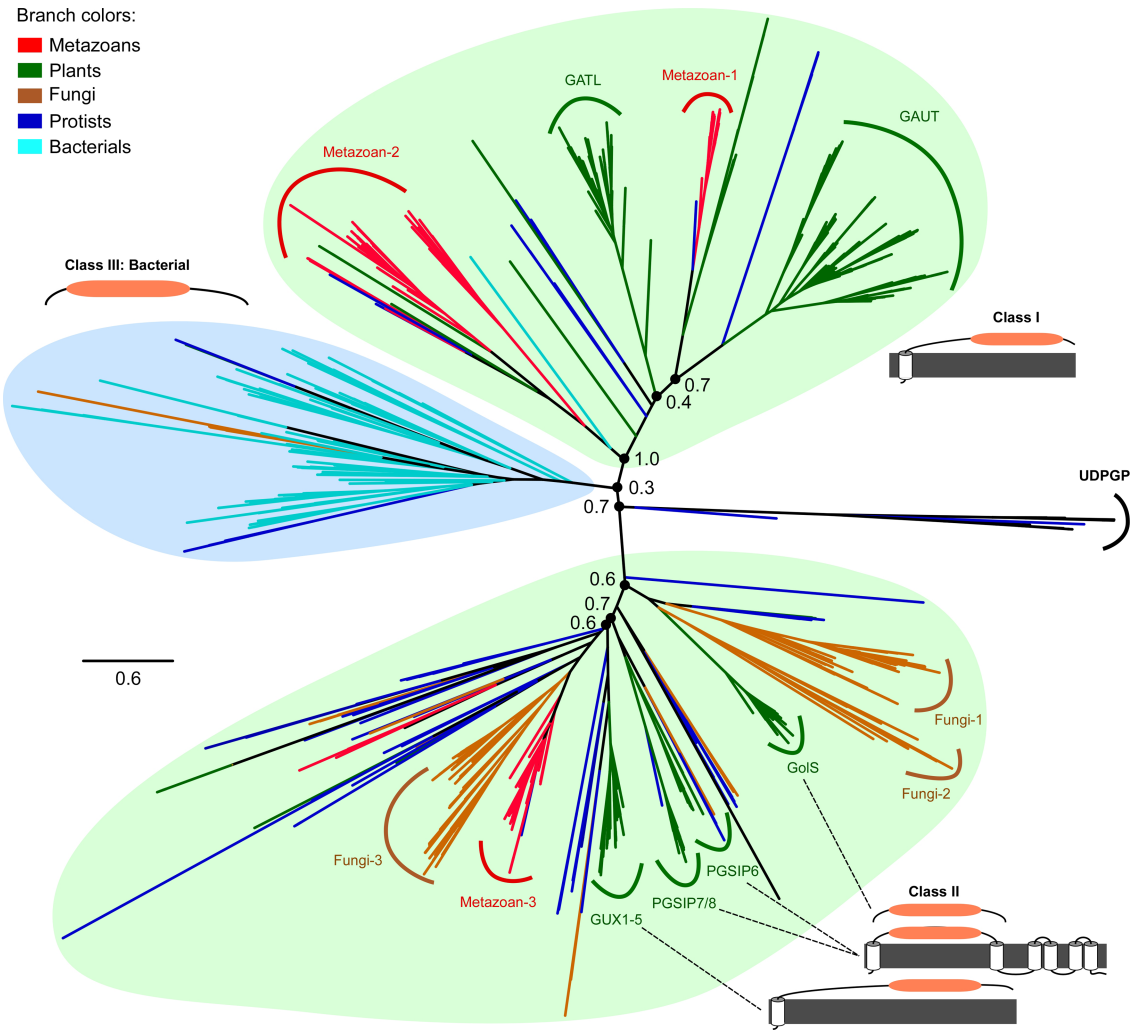


Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | The *moca1* mutant is hypersensitive to salt stress and *MOCA1* encodes a glucuronosyltransferase. a, **b**, Wild-type, *moca1* and complementation transgenic seedlings (*MOCA1 moca1*, endogenous *MOCA1* promoter-driven complementation lines; *MOCA1*^{ox} *moca1*, 35S-driven *MOCA1* overexpression complementation lines) have similar total amounts of remaining aequorin. The same seedlings used in Fig. 4b were treated with a solution containing 0.9 M $CaCl_2$ and 10% (v/v) ethanol to measure the total amount of remaining aequorin (**a**). Similar results were seen in ten separate experiments. The total amount of remaining aequorin in wild-type, *moca1* and transgenic plants was quantified from experiments as in **a** (**b**; mean ± s.d., $n = 40$ pools (30 seedlings per pool)). **c–e**, Complementation of salt hypersensitivity of growth in the *MOCA1 moca1* and *MOCA1*^{ox} *moca1* transgenic lines. Plants grown in medium in the absence (left) or presence of 60 mM NaCl (right) (**c**); and root length (**d**) and survival rate (**e**) from experiments in **c** (mean ± s.d.; $n = 20$ pools (16 seedlings per pool)). **f**, *MOCA1* gene expressed in stamens and flowerers. GUS expression was examined using transgenic plants carrying the *pMOCA1::GUS* construct as in Fig. 4d. Data are representative of more than ten independent experiments. **g**, Golgi membrane localization of both MOCA1 and mutant MOCA1 (mMOCA1) proteins. GFP fluorescence was analysed in seedlings expressing the *pMOCA1::MOCA1–GFP* construct (top) or the *pMOCA1::mMOCA1–GFP* construct (bottom). Data are representative of more than ten independent experiments. **h**, Diagram of plant GIPC structure adapted from previous studies[23,27,30]. Hydroxyl group and net negative charge as well as the outline of MOCA1 (IPUT1)-catalysed GIPC synthesis are illustrated. Hex, hexose; GlcA, glucuronic acid; Ins, inositol; VLCFA, very long chain fatty acid; LCB, long chain base.



Extended Data Fig. 9 | ITC analysis of binding of Na^+, K^+ and Li^+ to GIPCs (wild-type) and IPCs (*moca1*). **a**, **b**, ITC analysis of binding of Na^+ to IPCs. Lipid vesicles were produced from the *moca1* GIPC–IPC mixture as in Fig. 5b with >90% IPCs (Fig. 5d). The ITC data (**a**) and the plots of injected heat for automatic injections of NaCl solution into the sample cell containing the vesicle solution (**b**) are shown. Six independent ITC experiments were performed with similar results. **c**–**f**, The number of sites (**c**), heat change (ΔH) (**d**), entropy change (ΔS) (**e**), and binding constants (**f**) were derived from the fitted ITC from experiments similar to those in Fig. 6a, b for wild-type (GIPCs) and those in **a**, **b** for *moca1* (IPCs). Mean ± s.d., $n = 6$; ***$P < 0.001$; NS, not significant; **d**, $P = 0.787$; **e**, $P = 0.341$; **f**, $P = 0.407$). **g**, **h**, ITC analysis of binding of K^+ to GIPCs (wild-type) and IPCs (*moca1*). The number of sites (**g**) and binding constants (**h**) were derived from the fitted ITC from experiments similar to those for Na^+ binding to GIPCs and IPCs. Mean ± s.d., $n = 8$; **$P = 0.017$; NS, not significant; **h**, $P = 0.442$). The number of apparent K^+-binding sites for GIPCs and IPCs was 2.85 ± 0.35 and 2.42 ± 0.28, respectively ($P < 0.01$); K_d was 1.873 ± 0.823 mM and 1.338 ± 0.190 mM, respectively ($P > 0.05$). **i**, **j**, ITC analysis of binding of Li^+ to GIPCs (wild-type) and IPCs (*moca1*). The number of sites (**i**) and binding constants (**j**) were derived from the fitted ITC. Mean ± s.d., $n = 8$; ***$P < 0.001$; NS, not significant; **j**, $P = 0.878$). The number of apparent Li^+-binding sites for GIPCs and IPCs was 3.36 ± 0.18 and 2.93 ± 0.16, respectively ($P < 0.001$); K_d was 2.857 ± 1.273 mM and 2.873 ± 1.209 mM, respectively ($P > 0.05$). The K_d values for binding of the three monovalent cations to GIPCs are comparable to the estimated K_d for Na^+ in depolarizing ζ potentials (Fig. 5e), but much smaller than the estimated K_d for these cations in triggering increases in $[Ca^{2+}]_i$ (Figs. 1a, 3d–f), largely owing to the dilution of these cations by apoplastic water when the cations go through stomatal pores to reach the apoplastic side of the plasma membrane.



Extended Data Fig. 10 | Phylogeny of GT8 domains in fully sequenced eukaryotic and bacterial genomes. UTP-glucose-1-phosphate uridylyltransferase (UDPGP) domain-containing proteins in humans, yeast, and *Arabidopsis* were selected as outgroups to root the gene tree. UDPGP and GT8 belong to the same pfam clan GT_A (CL0110). Six groups of *Arabidopsis* GT8 genes, three groups of metazoan genes, and another three groups of fungal genes are identified and highlighted. Domain structures of plant GT8 genes are labelled similarly to these in Fig. 4a. The UDPGP domain-containing proteins are NP_006750, XP_005245519, and XP_006717380 in humans; NP_012889, NP_011851, and NP_010180.1 in yeast; and NP_197233, NP_186975, NP_181047, and NP_564372 in *Arabidopsis*.

nature research

Corresponding author(s): Zhangli Hu, Zhen-Ming Pei

Last updated by author(s): 2019/06/17

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	WinView/32 (Roper); Meta Morph 6.3, MetaMorph Basic Acquisition for Microscope and Meta Fluor (Molecular Devices);Origin software (Malvern Panalytical and OriginLab.
Data analysis	EXCEL16 software (Microsoft); Meta Morph 6.3, MetaMorph Basic Acquisition for Microscope and Meta Fluor (Molecular Devices); Image J; SAS 9.3 software (SAS Institute); Origin software (Malvern Panalytical and OriginLab).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Source Data related to Fig 1-5 and Extended Fig 2-6, 8, 9 are provided with manuscript. All lines and other data supporting the findings of this study are available from the corresponding author upon request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to predetermine sample size. The determined sample size was adequate as the differences between experimental groups was significant and reproducible.
Data exclusions	No data were excluded from the analyses.
Replication	The recordings and parameters were reproducible across multiple days and batches. All attempts of replication were successful.
Randomization	Randomization of samples were performed. Seedlings from different plates were collected.
Blinding	The investigators were not blinded to allocation during experiments and outcome assessment. In order to get as objective results as possible, in multiple experiments we had other researchers repeating the experiments.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

ARTICLE

https://doi.org/10.1038/s41586-019-1377-y

Modulation of cardiac ryanodine receptor 2 by calmodulin

Deshun Gong[1,5]*, Ximin Chi[1,5], Jinhong Wei[2,5], Gewei Zhou[1], Gaoxingyu Huang[1], Lin Zhang[2], Ruiwu Wang[2], Jianlin Lei[3], S. R. Wayne Chen[2]* & Nieng Yan[1,4]*

The high-conductance intracellular calcium (Ca^{2+}) channel RyR2 is essential for the coupling of excitation and contraction in cardiac muscle. Among various modulators, calmodulin (CaM) regulates RyR2 in a Ca^{2+}-dependent manner. Here we reveal the regulatory mechanism by which porcine RyR2 is modulated by human CaM through the structural determination of RyR2 under eight conditions. Apo-CaM and Ca^{2+}-CaM bind to distinct but overlapping sites in an elongated cleft formed by the handle, helical and central domains. The shift in CaM-binding sites on RyR2 is controlled by Ca^{2+} binding to CaM, rather than to RyR2. Ca^{2+}-CaM induces rotations and intradomain shifts of individual central domains, resulting in pore closure of the PCB95 and Ca^{2+}-activated channel. By contrast, the pore of the ATP, caffeine and Ca^{2+}-activated channel remains open in the presence of Ca^{2+}-CaM, which suggests that Ca^{2+}-CaM is one of the many competing modulators of RyR2 gating.

Cardiac muscle contraction is triggered by Ca^{2+} flux into the cytosol, initially from the extracellular environment, mediated by $Ca_v1.2$, and subsequently from the sarcoplasmic reticulum Ca^{2+} store, mediated by RyR2[1–3]. Ryanodine receptors are the largest known ion channels and consist of a homotetramer with a molecular mass of more than 2 megadaltons. More than 80% of the protein folds into a multi-domain cytoplasmic assembly that senses interactions with a variety of modulators, which range from ions to proteins[4–6]. The precise regulation of RyR2 activity is critical for each heartbeat. Aberrant activity of RyR2 is associated with life-threatening cardiac arrhythmias[7–10].

The 17-kDa protein CaM is an essential calcium sensor that has a central role in most calcium signalling events[11]. CaM consists of roughly symmetrical N- and C-terminal lobes (N- and C-lobes hereafter), joined by a flexible hinge[12,13]. Each lobe can cooperatively bind to two Ca^{2+} ions, with a micromolar-range binding affinity, via two EF-hand (helices E and F-hand) motifs. Upon Ca^{2+} binding, the exposure of several hydrophobic residues in both lobes facilitates CaM binding to the target sequence. CaM interacts directly with ryanodine receptors with a 1:1 stoichiometry of the CaM–RyR protomers[14,15] and binding affinity at nanomolar range[14].

Regulation of ryanodine receptors by CaM, however, is isoform-specific. CaM shows biphasic regulation of RyR1, acting as a weak activator at nanomolar levels of Ca^{2+} (apo-CaM) and an inhibitor at micromolar levels of Ca^{2+} (Ca^{2+}-CaM)[14,16]. By contrast, apo-CaM has no effect[17] or an inhibitory effect on RyR2[14], whereas Ca^{2+}-CaM inhibits RyR2[14]. CaM has also been shown to facilitate the termination of store-overload-induced Ca^{2+} release (SOICR)[18]. Aberrant interactions between CaM and RyR2 are associated with heart failure[19–22], and correction of impaired CaM–RyR2 interactions may serve as a therapy for lethal arrhythmia in pressure-overload-induced heart failure[23].

Structural characterization of RyR–CaM complexes has been limited to low-resolution electron microscopy maps that suggest two overlapping, but distinct, binding sites in RyR1 for apo- and Ca^{2+}-CaM[24–26]. A peptide that corresponds to residues 3614–3643 of RyR1 (residues 3581–3612 in the central domain of RyR2) binds to both apo- and Ca^{2+}-CaM[15,27]. The crystal structure of Ca^{2+}-CaM bound to the peptide revealed hydrophobic anchors in the N and C termini of the peptide that accommodate the C- and N-lobes of Ca^{2+}-CaM, respectively[28].

To elucidate the modulation of RyR2 by CaM, we report eight cryo-electron microscopy (cryo-EM) structures of RyR2 that collectively reveal molecular recognition characteristics for different forms of CaM and provide insights into the regulation of RyR2 channel gating by CaM.

Structures of RyR2 under eight conditions

To achieve a better understanding of RyR2 modulation by CaM (Extended Data Fig. 1a, b), we determined the cryo-EM structures of the porcine RyR2 (Extended Data Fig. 1c) under the following eight conditions.

Condition (1) consisted of RyR2 bound to FKBP12.6 and apo-CaM (hereafter FKBP12.6/apo-CaM) and was used to assess the apo-CaM binding site. Condition (2) consisted of RyR2 bound to FKBP12.6 and a Ca^{2+}-binding-deficient CaM mutant that mimics apo-CaM[24,29] (CaM-M) in the presence of ATP, caffeine and low [Ca^{2+}] (hereafter FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/CaM-M), this structure was used to investigate the mechanism for the binding-location switch of CaM. Condition (3) consisted of RyR2 bound to FKBP12.6 in the presence of ATP, caffeine and low [Ca^{2+}] (FKBP12.6/ATP/caffeine/low-[Ca^{2+}]), the presence of which maximizes the open state. Condition (4) consisted of RyR2 bound to FKBP12.6 and Ca^{2+}-CaM in the presence of ATP, caffeine and low [Ca^{2+}] (hereafter FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/Ca^{2+}-CaM); this condition was used to examine the effect of Ca^{2+}-CaM on the open RyR2 channel in the presence of FKBP12.6, ATP, caffeine and low [Ca^{2+}]. Conditions (5) and (6) corresponded to conditions (3) and (4), respectively, but were treated with CHAPS and DOPC instead of digitonin. Condition (7) consisted of RyR2 in high [Ca^{2+}] in the presence of FKBP12.6, ATP, caffeine and Ca^{2+}-CaM

[1]Beijing Advanced Innovation Center for Structural Biology, Tsinghua-Peking Joint Center for Life Sciences, School of Life Sciences, Tsinghua University, Beijing, China. [2]Libin Cardiovascular Institute of Alberta, Department of Physiology and Pharmacology, University of Calgary, Calgary, Alberta, Canada. [3]Technology Center for Protein Sciences, Ministry of Education Key Laboratory of Protein Sciences, School of Life Sciences, Tsinghua University, Beijing, China. [4]Present address: Department of Molecular Biology, Princeton University, Princeton, NJ, USA. [5]These authors contributed equally: Deshun Gong, Ximin Chi, Jinhong Wei. *e-mail: gds13@tsinghua.org.cn; swchen@ucalgary.ca; nyan@princeton.edu

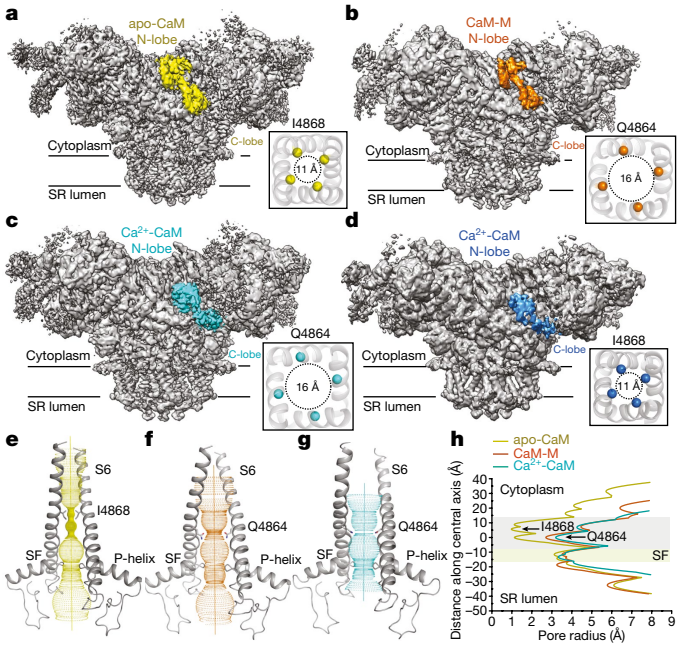


Fig. 1 | Cryo-EM structures of the RyR2–CaM complexes. a–d, Overall electron microscopy maps for four indicated complexes. **a**, RyR2 in the presence of FKBP12.6/apo-CaM (3.6 Å). **b**, RyR2 in the presence of FKBP12.6/ATP/caffeine/low [Ca^{2+}]/CaM-M (4.2 Å). **c**, RyR2 in the presence of FKBP12.6/ATP/caffeine/high [Ca^{2+}]/Ca^{2+}-CaM (3.9 Å). **d**, RyR2 in the presence of PCB95/low-[Ca^{2+}]/Ca^{2+}-CaM (4.4 Å). Insets, cytoplasmic views of the channel gates. The dashed circles indicate the distances between the C_α atoms of the gating residues in the diagonal protomers. See Extended Data Table 1 for details of the structures. SR, sarcoplasmic reticulum. **e**, The pore of RyR2 remains closed in the FKBP12.6/apo-CaM structure. The ion permeation path, calculated by HOLE[41], is illustrated as yellow dots. SF, selectivity filter. **f**, **g**, Open pores of RyR2 in FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/CaM-M (**f**) and FKBP12.6/ATP/caffeine/high-[Ca^{2+}]/Ca^{2+}-CaM (**g**). **h**, The corresponding pore radii of RyR2 for the three structures shown in **e**–**g**. Electron microscopy maps were generated in Chimera and contoured at levels of 0.027, 0.022, 0.02 and 0.015 for **a**–**d**, respectively. All structures were prepared using PyMOL (http://www.pymol.org).

(hereafter FKBP12.6/ATP/caffeine/high-[Ca^{2+}]/Ca^{2+}-CaM); this environment was used to achieve a better resolution for Ca^{2+}-CaM. Condition (8) consisted of RyR2 bound to Ca^{2+}-CaM in the presence of 2,2′,3,5′,6-pentachlorobiphenyl (PCB95) and low [Ca^{2+}] (hereafter PCB95/low-[Ca^{2+}]/Ca^{2+}-CaM); this condition was used to investigate the effect of Ca^{2+}-CaM on PCB95 and the Ca^{2+}-activated RyR2 channel[30]. The eight conditions and corresponding structures are summarized in Supplementary Table 1 and Extended Data Table 1, respectively.

All cryo-EM datasets were processed following the same procedure (Extended Data Fig. 2). The FKBP12.6/apo-CaM RyR2 structure was determined at an overall resolution of 3.6 Å, the highest among all of the available RyR2 structures (Fig. 1a and Extended Data Figs. 1i, 3a, 4a–i). The secondary structural elements of apo-CaM were clearly resolved (Fig. 1a and Extended Data Fig. 3b). The FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/CaM-M RyR2 structure was determined at an overall resolution of 4.2 Å, in which the well-resolved CaM-M is positioned similarly to apo-CaM on RyR2 (Fig. 1b and Extended Data Figs. 1i, 3c, d). The densities for both lobes of Ca^{2+}-CaM are visible in the FKBP12.6/ATP/caffeine/high-[Ca^{2+}]/Ca^{2+}-CaM RyR2 structure (3.9 Å resolution) and PCB95/low-[Ca^{2+}]/Ca^{2+}-CaM RyR2 structure (4.4 Å resolution) in which the N-lobe is better-resolved than the C-lobe. By contrast, only one lobe of Ca^{2+}-CaM—the N lobe as judged from the comparison with the structure of FKBP12.6/ATP/caffeine/high-[Ca^{2+}]/Ca^{2+}-CaM at 3.9 Å—is discernible in the FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/Ca^{2+}-CaM structure with a resolution of 4.2 Å (Fig. 1c, d and Extended Data Figs. 1, 3).

With regard to the gating state of the eight structures reported here, the clearly resolved Ile4868 residues on the S6 helical bundle of FKBP12.6/apo-CaM constitute the constriction site with a radius of approximately 1 Å, which is identical to the previously reported apo-RyR2 in the closed state[30] (Fig. 1e, h). The constriction site appears to shift to Gln4864 with an expanded radius of around 3 Å in conditions (2), (3) and (5)–(7) (Fig. 1f–h and Extended Data Fig. 1d, h), similar to that in the open PCB95/low-[Ca^{2+}] structure[30]. However, the density for the side chain of Gln4864 in FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/Ca^{2+}-CaM is not well-resolved. We, therefore, compared the distances of C_α atoms of the gating residues in the diagonal protomers, which are approximately 16 Å for FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/Ca^{2+}-CaM and 11 Å for FKBP12.6/apo-CaM (Fig. 1a and Extended Data Fig. 1e). The constriction site in PCB95/low-[Ca^{2+}]/Ca^{2+}-CaM is constituted by Ile4868, for which the diagonal distance of the C_α atoms is approximately 11 Å—similar to that in FKBP12.6/apo-CaM (Fig. 1a, d).

As seen in RyR1, the Ca^{2+}-, ATP- and caffeine-binding sites are located at the interfaces between the central and channel domains of RyR2[31] (Extended Data Fig. 4j–m).

Location of apo-CaM in RyR2

Consistent with the low-resolution structure of RyR1, apo-CaM is located in an elongated cleft formed by the handle, helical and central domains of RyR2 in FKBP12.6/apo-CaM[24,25] (Fig. 2a). The N-lobe is stuck in the upper half of the cleft formed by helical domain 1 (HD1), whereas the C-lobe is located at the bottom edge of the cleft surrounded by the handle and central domains of RyR2 (Fig. 2a).

The FKBP12.6/apo-CaM structure reveals five surface patches on RyR2 that interact with apo-CaM. The N-lobe interacts with RyR2 through three interfaces that are mainly located in HD1 (Fig. 2b and Extended Data Fig. 5a). The most prominent interface is formed between the N terminus of helix 4 (N4) in the N-lobe and the C termini of helices 2b and α1 in HD1, and is mainly mediated by extensive hydrophobic residues. Phe66, Pro67 and Leu70 on N4 probably interact with Tyr2203 in helix 2b and Tyr2157 (human Tyr2156) in helix α1. Ile10 in N1 may also interact with Tyr2157 (Extended Data Fig. 5a, g). The human Y2156C variant is linked to catecholaminergic polymorphic ventricular tachycardia[8]. The second interface is mediated by charged residues between the N terminus of N1 in the N-lobe and helix α1 in HD1 and the N terminus of helix α0 in the central domain. The third interface is formed between a region rich in acidic residues in N3 of the N-lobe and Lys2558 in helix 8b of HD1 (Fig. 2b and Extended Data Fig. 5a).

The C-lobe interacts with RyR2 through two interfaces (Fig. 2c). One is consistent with previous reports[15,27]. Residues 3593–3607 in the central domain folds into a newly resolved helix α 'minus 1' (helix α−1) that is enclosed by the hydrophobic cavity of the C-lobe, representing the primary interface. Phe3604 serves as a hydrophobic anchor for the hydrophobic cavity of the C-lobe. A minor interface is formed between helix 12 and the C terminus of helix 11 in the handle domain with the C terminus of C1 and the loop between C2 and C3 in the C-lobe, also through hydrophobic interactions (Fig. 2c and Extended Data Fig. 5b, c, h).

Shift of CaM-binding site in RyR2 after Ca^{2+} loading

CaM markedly slips down along the cleft after Ca^{2+} loading, making extensive interactions with the central domain (Fig. 2d). The N-lobe is anchored by the central domain and the C-lobe drops beyond the cleft, coordinated only by helix α−1 (Fig. 2d). The limited contact may explain the structural flexibility of the C-lobe.

The binding of Ca^{2+}-CaM with intact RyR2 is similar to that of Ca^{2+}-CaM with the RyR1 peptide[28]. The N- and C-lobes of CaM interact with the C- and N termini of helix α−1, respectively (Fig. 2d). Phe3604 and Trp3588 anchor the hydrophobic cavities of the N- and C-lobes, respectively (Extended Data Fig. 5d–f, i). An additional interface is formed between the N terminus of N3 and the C terminus of helix α9 in the central domain to further stabilize the binding of N-lobe. Asp51

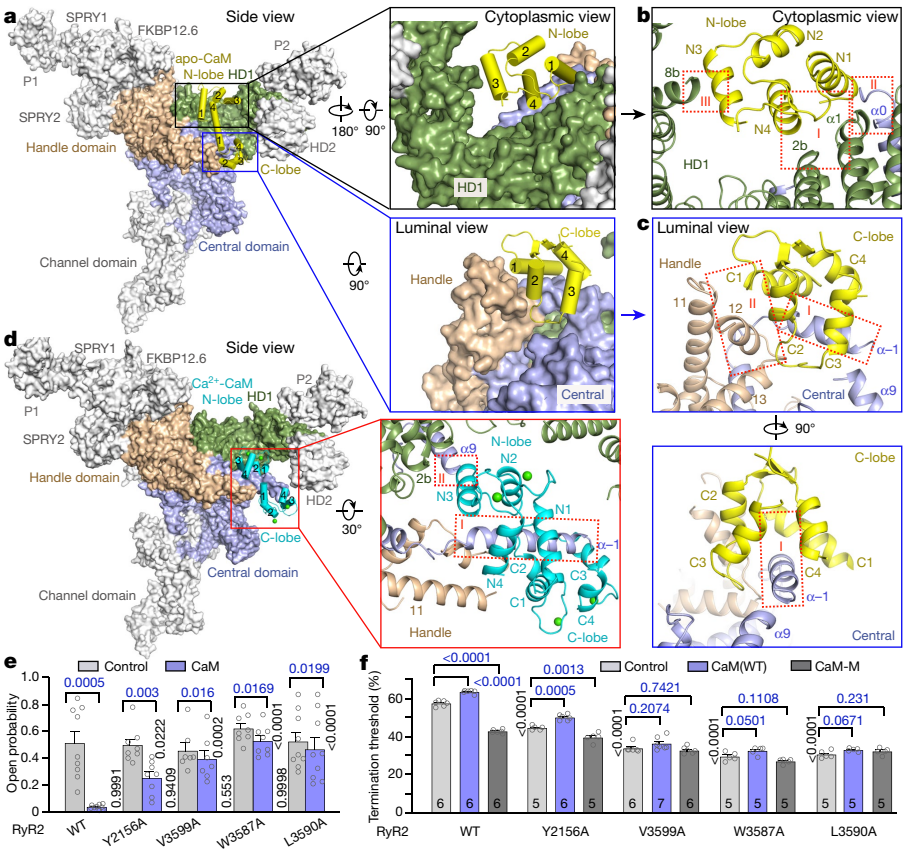


Fig. 2 | Interfaces between CaM and RyR2. a, Apo-CaM is located in a cleft formed by the handle, helical and central domains of RyR2. One RyR2 protomer is shown in domain-coloured surface view. **b**, **c**, Multiple interfaces between RyR2 and apo-CaM. The red, dashed boxes indicate the interfaces. **d**, A previously unresolved helix α−1 on RyR2 serves as the primary docking site for both lobes of Ca^{2+}-CaM. **e**, Functional validation of the observed interfaces between mouse RyR2 and human CaM. Open probabilities of single RyR2 channels before (control) and after addition of CaM(WT) (1 μM). Data are mean ± s.e.m. from RyR2(WT) ($n = 9$), RyR2(Y2156A) ($n = 8$), RyR2(V3599A) ($n = 7$), RyR2(W3587A) ($n = 8$) and RyR2(L3590A) ($n = 9$ single channels) and analysed by paired, two-sided Student's *t*-test (versus its own control) with *P* values shown in blue and by one-way analysis of variance (ANOVA) with a Dunnett's post hoc test (versus RyR2(WT) control and RyR2(WT) with CaM(WT), respectively) with adjusted *P* values shown in black. **f**, The termination threshold of Ca^{2+} release in $Ryr2^{WT}$- and *Ryr2*-mutant-expressing HEK293 cells transfected with no CaM (control), CaM(WT) or CaM-M. Data are mean ± s.e.m. with the number of independent experiments for each condition shown and analysed by one-way ANOVA with a Dunnett's post hoc test with adjusted *P* values shown in blue (versus its own control) and in black (versus RyR2(WT) control).

in N3 probably interacts with Arg2209 in helix 2b of HD1 (Fig. 2d and Extended Data Fig. 5e).

Functional validation of RyR2–CaM interfaces

We assessed the effect of mutations in RyR2 and CaM that are located in the structurally revealed interfaces on CaM regulation of RyR2. Wild-type CaM (CaM(WT)) strongly reduced the open probability of single wild-type mouse RyR2 (RyR2(WT)) channels. The following mutations markedly reduced CaM inhibition of single RyR2 channels: RyR2(Y2156A) (porcine Y2157A) near an interface between the apo-CaM N-lobe and RyR2, RyR2(V3599A) near both the apo-CaM C-lobe–RyR2 and Ca^{2+}-CaM N-lobe–RyR2 interfaces, and RyR2(W3587A) and RyR2(L3590A) near the Ca^{2+}-CaM C-lobe–RyR2 interface (Fig. 2e and Extended Data Fig. 6).

We next examined the effect of these RyR2 mutations on the termination of RyR2-mediated SOICR. As previously shown[18], CaM(WT) increased Ca^{2+} release termination in RyR2(WT)-expressing HEK293 cells, whereas CaM-M reduced Ca^{2+} release termination (that is, a longer calcium release) (Fig. 2f and Extended Data Fig. 7a–c). Consistent with their effect on single RyR2 channels, all four mutations in RyR2 significantly reduced Ca^{2+} release termination in HEK293 cells, probably by impairing the effect of endogenous CaM on RyR2 inhibition (Fig. 2f). These RyR2 mutations also reduced or abolished the effect of exogenously expressed CaM(WT) and CaM-M on Ca^{2+} release termination (Fig. 2f), but had little or no effect on SOICR activation or store capacity (Extended Data Fig. 7d, e).

We also assessed the functional importance of CaM residues near the RyR2–CaM interfaces. Mutations in CaM near the apo-CaM N-lobe–RyR2(K2153/Y2156) interface (CaM(E15A), CaM(F66A) and CaM(L70A)), near the apo-CaM C-lobe–RyR2(V3599) interface (CaM(M110A) and CaM(F142A)), near the Ca^{2+}-CaM N-lobe–RyR2(V3599) interface (CaM(F20A) and CaM(F69A)) and near the Ca^{2+}-CaM C-lobe–RyR2(W3587/L3590) interface (CaM(F93A), CaM(L106A) and CaM(M146A))—as with CaM-M—significantly reduced the effect of CaM on Ca^{2+} release termination compared to CaM(WT) (Extended Data Fig. 7f). All CaM mutations except for CaM(F20A) and CaM(F142A) had little or no effect on SOICR activation or store capacity (Extended Data Fig. 7g, h). Note that some mutations in CaM may induce conformational changes, thus affecting CaM–RyR2 interactions allosterically. Collectively, these functional studies support the importance of the newly identified RyR2–CaM interfaces in CaM regulation of RyR2.

Ca^{2+}-dependent shift in CaM-binding sites on RyR2

The location and conformation of CaM-M and apo-CaM are identical in the structures (Fig. 3a), which suggests that the positional switch for apo-CaM and Ca^{2+}-CaM results from the distinct conformations of CaM after Ca^{2+} loading instead of a direct effect of Ca^{2+} on RyR2. The CaM lobes have previously been reported in three

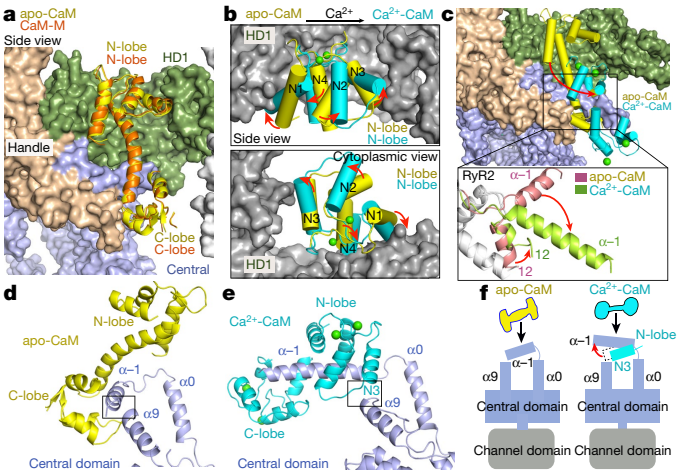


Fig. 3 | Molecular basis for the shift of binding site for CaM after Ca^{2+} loading. a, Ca^{2+} binding to RyR2 is not responsible for the positional shift in apo-CaM and Ca^{2+}-CaM. In the presence of 20 μM Ca^{2+}, the binding site for CaM-M remains the same as for apo-CaM. The two structures are superimposed relative to CaM. **b**, Upon Ca^{2+} loading, the expansion of CaM structure may lead to steric hindrance between the N-lobe and HD1. The N-lobes are superimposed. Red arrows indicate directions of conformational changes from apo-CaM to Ca^{2+}-CaM. **c**, The shift in the CaM-binding site is accompanied by marked conformational changes in helices α−1 and 12 in RyR2. Red arrows indicate the conformational changes in RyR2 from FKBP12.6/apo-CaM to FKBP12.6/ATP/caffeine/high-[Ca^{2+}]/Ca^{2+}-CaM. **d**, In FKBP12.6/apo-CaM, helix α−1 contacts helix α9. **e**, In FKBP12.6/ATP/caffeine/high-[Ca^{2+}]/Ca^{2+}-CaM, helix α−1 of RyR2 is positioned away from helix α9. Helix N3 in the N-lobe contacts helix α9 instead. **f**, Helix α−1 of RyR2 serves as an essential anchor for CaM. Red arrow indicates the direction of conformational change of helix α−1 from FKBP12.6/apo-CaM to FKBP12.6/ATP/caffeine/high-[Ca^{2+}]/Ca^{2+}-CaM.

conformations—open, semi-open and closed—when interacting with other proteins[32–34]. Comparative docking analysis of these conformations into the electron microscopy reconstruction for FKBP12.6/ATP/caffeine/high-[Ca^{2+}]/Ca^{2+}-CaM suggests an open conformation for the N-lobe (Extended Data Fig. 8a, b). An important distinction between the open and semi-open or closed C-lobes is the slightly larger helical angle between C1 and C4, both of which were resolved in the map and conformed to the open state (Extended Data Fig. 8c, d).

These analyses are consistent with the observations in the crystal structure of RyR1 peptide-bound Ca^{2+}-CaM[28]. Upon Ca^{2+} loading, the compact structure of apo-CaM is relaxed to Ca^{2+}-CaM[35], which slips down towards the N-terminal part of helix α−1, consistent with previous studies[36].

Helix α−1 and helix 12 are two helices that were only resolved in the structures with CaM, probably owing to stabilization of these segments through CaM binding. Pronounced shifts in these two helices are observed between the structures bound to apo-CaM and Ca^{2+}-CaM. In the presence of Ca^{2+}-CaM, helix α−1 swings by around 90° and the N-terminal half of helix 12 also bends by nearly 90° (Fig. 3c). Accordingly, helix α−1, which contacts helix α9 of the central domain in FKBP12.6/apo-CaM, is positioned away from helix α9 in FKBP12.6/ATP/caffeine/high-[Ca^{2+}]/Ca^{2+}-CaM. Now helix α9 is in contact with helix 3 in the N-lobe of Ca^{2+}-CaM (Fig. 3d–f). The CaM-bound RyR2 structures shown here reveal that helix α−1 serves as an essential anchor for CaM (Fig. 3f).

Inhibitory modulation of RyR2 by Ca^{2+}-CaM

Both caffeine and ATP are located at the interfaces between the U-motif and O-ring, locking them into a stable unit that stabilizes the open state (Extended Data Fig. 9a). Caffeine and ATP counteract inhibitory effect of Ca^{2+}-CaM on RyR2, manifested by the lack of intradomain change of individual central domains. The pore remains open (Extended Data Figs. 1e, 9b). Nevertheless, Ca^{2+}-CaM induces an anticlockwise rotation of the central domains in the cytoplasmic view in the same direction as that from the open to the closed state[30,37] (Extended Data Fig. 9c and Supplementary Video 1). The central domains undergo similar shifts in the presence of high concentrations of Ca^{2+} and CaM, but not in the CHAPS plus DOPC condition (Extended Data Fig. 10a, b).

By contrast, RyR2 activated by PCB95 and Ca^{2+} is closed after addition of Ca^{2+}-CaM (Fig. 1d). Detailed structural examination shows an anticlockwise rotation of the central domains and outward motions of the auxiliary motifs of the individual central domain, including helices α0, α1, α4 and the U-motif, with respect to the centre of the concave surface (Extended Data Figs. 9d, 10c). The motion of the U-motif appears to release the pulling force for the dilation of the S6 helix, resulting in closure of the pore (Extended Data Fig. 9e). Taken together, these results indicate that the inhibitory force of Ca^{2+}-CaM is sufficient to overcome the synergistic activation of RyR2 by PCB95 and Ca^{2+}, but not by the collective effect of ATP, caffeine and Ca^{2+} (Fig. 4).

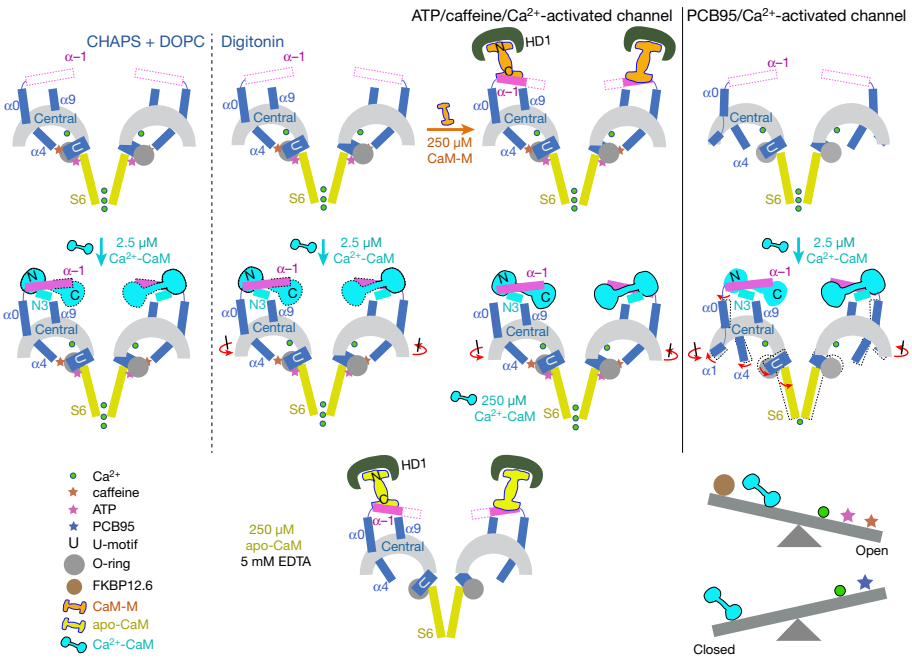


Fig. 4 | Schematic of RyR2 modulation by CaM. The two RyR2 structures on the left were obtained in the presence of CHAPS and DOPC instead of digitonin, which was used for all other structural determinations. Despite a rotation (indicated by red arrows) of the central domains, the pore of ATP, caffeine and Ca^{2+}-activated RyR2 channel remains open in the presence of Ca^{2+}-CaM under these two conditions (left four). By contrast, Ca^{2+}-CaM leads to closure of PCB95 and Ca^{2+}-activated RyR2 channel (right two). Our study demonstrates that the inhibitory force of Ca^{2+}-CaM can overcome the synergistic activation by PCB95 and Ca^{2+} but not by ATP, caffeine and Ca^{2+} (bottom right). The top right structure PCB95/low-Ca^{2+} was obtained from a previous study (RCSB Protein Data Bank (PDB) code 5GOA).

Discussion

Among the reported RyR2 sequences for CaM-binding[38], residues 4247–4277 are invisible in our structure. The segment (residues 2023–2039) that contains the protein kinase A phosphorylation site Ser2032 is resolved in our structures, but shows no interaction with CaM even at high concentrations[39]. The other two sequences (1942–1966 and 3582–3608) bind to CaM in our structures (Extended Data Fig. 10d). Despite the distinct effects of apo-CaM on RyR1 and RyR2[14], the primary apo-CaM-binding sequences are invariant in these two isoforms (Extended Data Fig. 10e). The molecular determinants for the functional difference are yet to be revealed.

The inhibitory mechanisms by which CaM regulates RyR2 must be investigated on an already opened channel by structural biology. As the presence of micromolar-range Ca^{2+} is required for opening ryanodine receptors by cryo-EM studies[30–32,37], it is impractical to obtain an open structure in the absence of micromolar concentrations of Ca^{2+} (apo-CaM form). Although the location and conformation of CaM-M appear identical to those of apo-CaM, it has previously been reported that modulation of RyR2 by CaM-M is distinct from that by apo-CaM[40]—although the mechanism needs to be investigated further.

Owing to extensive interactions between the U-motif and O-ring, the two undergo coupled motions during channel gating[30,37]. The presence of caffeine and ATP locks them into a more rigid structure, probably increasing the energy barrier for inhibiting RyR2 by Ca^{2+}-CaM. By contrast, the PCB95- and Ca^{2+}-activated RyR2 channel can be effectively closed by Ca^{2+}-CaM. Therefore, the gating state of RyR2 is defined by the combined effect of competing stimulatory and inhibitory regulators (Fig. 4). It remains to be investigated whether the conclusions presented here can be recapitulated for other ryanodine receptor isoforms or in lipid bilayers and the relevance to disease-related mutations (Extended Data Fig. 10f, g).

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1377-y.

Received: 26 January 2019; Accepted: 4 June 2019;
Published online 5 July 2019.

1. Fabiato, A. Calcium-induced release of calcium from the cardiac sarcoplasmic reticulum. *Am. J. Physiol.* **245**, C1–C14 (1983).
2. Nakai, J. et al. Primary structure and functional expression from cDNA of the cardiac ryanodine receptor/calcium release channel. *FEBS Lett.* **271**, 169–177 (1990).
3. Otsu, K. et al. Molecular cloning of cDNA encoding the Ca^{2+} release channel (ryanodine receptor) of rabbit cardiac muscle sarcoplasmic reticulum. *J. Biol. Chem.* **265**, 13472–13483 (1990).
4. Rodney, G. G., Williams, B. Y., Strasburg, G. M., Beckingham, K. & Hamilton, S. L. Regulation of RYR1 activity by Ca^{2+} and calmodulin. *Biochemistry* **39**, 7807–7812 (2000).
5. Timerman, A. P. et al. The ryanodine receptor from canine heart sarcoplasmic reticulum is associated with a novel FK-506 binding protein. *Biochem. Biophys. Res. Commun.* **198**, 701–706 (1994).
6. Yamaguchi, N., Xu, L., Pasek, D. A., Evans, K. E. & Meissner, G. Molecular basis of calmodulin binding to cardiac muscle Ca^{2+} release channel (ryanodine receptor). *J. Biol. Chem.* **278**, 23480–23486 (2003).
7. Laitinen, P. J. et al. Mutations of the cardiac ryanodine receptor (RyR2) gene in familial polymorphic ventricular tachycardia. *Circulation* **103**, 485–490 (2001).
8. Medeiros-Domingo, A. et al. The *RYR2*-encoded ryanodine receptor/calcium release channel in patients diagnosed previously with either catecholaminergic polymorphic ventricular tachycardia or genotype negative, exercise-induced long QT syndrome: a comprehensive open reading frame mutational analysis. *J. Am. Coll. Cardiol.* **54**, 2065–2074 (2009).
9. Priori, S. G. & Chen, S. R. Inherited dysfunction of sarcoplasmic reticulum Ca^{2+} handling and arrhythmogenesis. *Circ. Res.* **108**, 871–883 (2011).
10. Priori, S. G. et al. Mutations in the cardiac ryanodine receptor gene (*hRyR2*) underlie catecholaminergic polymorphic ventricular tachycardia. *Circulation* **103**, 196–200 (2001).
11. Hoeflich, K. P. & Ikura, M. Calmodulin in action: diversity in target recognition and activation mechanisms. *Cell* **108**, 739–742 (2002).
12. Babu, Y. S. et al. Three-dimensional structure of calmodulin. *Nature* **315**, 37–40 (1985).
13. Copley, R. R., Schultz, J., Ponting, C. P. & Bork, P. Protein families in multicellular organisms. *Curr. Opin. Struct. Biol.* **9**, 408–415 (1999).
14. Balshaw, D. M., Xu, L., Yamaguchi, N., Pasek, D. A. & Meissner, G. Calmodulin binding and inhibition of cardiac muscle calcium release channel (ryanodine receptor). *J. Biol. Chem.* **276**, 20144–20153 (2001).
15. Moore, C. P. et al. Apocalmodulin and Ca^{2+} calmodulin bind to the same region on the skeletal muscle Ca^{2+} release channel. *Biochemistry* **38**, 8532–8537 (1999).
16. Tripathy, A., Xu, L., Mann, G. & Meissner, G. Calmodulin activation and inhibition of skeletal muscle Ca^{2+} release channel (ryanodine receptor). *Biophys. J.* **69**, 106–119 (1995).
17. Fruen, B. R., Bardy, J. M., Byrem, T. M., Strasburg, G. M. & Louis, C. F. Differential Ca^{2+} sensitivity of skeletal and cardiac muscle ryanodine receptors in the presence of calmodulin. *Am. J. Physiol. Cell Physiol.* **279**, C724–C733 (2000).
18. Tian, X., Tang, Y., Liu, Y., Wang, R. & Chen, S. R. Calmodulin modulates the termination threshold for cardiac ryanodine receptor-mediated Ca^{2+} release. *Biochem. J.* **455**, 367–375 (2013).
19. Hino, A. et al. Enhanced binding of calmodulin to the ryanodine receptor corrects contractile dysfunction in failing hearts. *Cardiovasc. Res.* **96**, 433–443 (2012).
20. Lavorato, M. et al. Dyad content is reduced in cardiac myocytes of mice with impaired calmodulin regulation of RyR2. *J. Muscle Res. Cell Motil.* **36**, 205–214 (2015).
21. Yamaguchi, N. et al. Cardiac hypertrophy associated with impaired regulation of cardiac ryanodine receptor by calmodulin and S100A1. *Am. J. Physiol. Heart Circ. Physiol.* **305**, H86–H94 (2013).
22. Yamaguchi, N., Takahashi, N., Xu, L., Smithies, O. & Meissner, G. Early cardiac hypertrophy in mice with impaired calmodulin regulation of cardiac muscle Ca release channel. *J. Clin. Invest.* **117**, 1344–1353 (2007).
23. Kato, T. et al. Correction of impaired calmodulin binding to RyR2 as a novel therapy for lethal arrhythmia in the pressure-overloaded heart failure. *Heart Rhythm* **14**, 120–127 (2017).
24. Huang, X., Fruen, B., Farrington, D. T., Wagenknecht, T. & Liu, Z. Calmodulin-binding locations on the skeletal and cardiac ryanodine receptors. *J. Biol. Chem.* **287**, 30328–30335 (2012).
25. Samsó, M. & Wagenknecht, T. Apocalmodulin and Ca^{2+}-calmodulin bind to neighboring locations on the ryanodine receptor. *J. Biol. Chem.* **277**, 1349–1353 (2002).
26. Wagenknecht, T. et al. Locations of calmodulin and FK506-binding protein on the three-dimensional architecture of the skeletal muscle ryanodine receptor. *J. Biol. Chem.* **272**, 32463–32471 (1997).
27. Yamaguchi, N., Xin, C. & Meissner, G. Identification of apocalmodulin and Ca^{2+}-calmodulin regulatory domain in skeletal muscle Ca^{2+} release channel, ryanodine receptor. *J. Biol. Chem.* **276**, 22579–22585 (2001).
28. Maximciuc, A. A., Putkey, J. A., Shamoo, Y. & Mackenzie, K. R. Complex of calmodulin with a ryanodine receptor target reveals a novel, flexible binding mode. *Structure* **14**, 1547–1556 (2006).
29. Maune, J. F., Klee, C. B. & Beckingham, K. Ca^{2+} binding and conformational change in two series of point mutations to the individual Ca^{2+}-binding sites of calmodulin. *J. Biol. Chem.* **267**, 5286–5295 (1992).
30. Peng, W. et al. Structural basis for the gating mechanism of the type 2 ryanodine receptor RyR2. *Science* **354**, aah5324 (2016).
31. des Georges, A. et al. Structural basis for gating and activation of RyR1. *Cell* **167**, 145–157 (2016).
32. Wei, R. et al. Structural insights into Ca^{2+}-activated long-range allosteric channel gating of RyR1. *Cell Res.* **26**, 977–994 (2016).
33. Wang, C. et al. Structural analyses of Ca^{2+}/CaM interaction with NaV channel C-termini reveal mechanisms of calcium-dependent regulation. *Nat. Commun.* **5**, 4896 (2014).
34. Wang, C., Chung, B. C., Yan, H., Lee, S. Y. & Pitt, G. S. Crystal structure of the ternary complex of a NaV C-terminal domain, a fibroblast growth factor homologous factor, and calmodulin. *Structure* **20**, 1167–1176 (2012).
35. Jurado, L. A., Chockalingam, P. S. & Jarrett, H. W. Apocalmodulin. *Physiol. Rev.* **79**, 661–682 (1999).
36. Rodney, G. G. et al. Calcium binding to calmodulin leads to an N-terminal shift in its binding site on the ryanodine receptor. *J. Biol. Chem.* **276**, 2069–2074 (2001).
37. Bai, X. C., Yan, Z., Wu, J., Li, Z. & Yan, N. The central domain of RyR1 is the transducer for long-range allosteric gating of channel opening. *Cell Res.* **26**, 995–1006 (2016).
38. Brohus, M., Søndergaard, M. T., Chen, S. R. W., van Petegem, F. & Overgaard, M. T. Ca^{2+}-dependent calmodulin binding to cardiac ryanodine receptor (RyR2) calmodulin-binding domains. *Biochem. J.* **476**, 193–209 (2019).
39. Xiao, B. et al. Characterization of a novel PKA phosphorylation site, serine-2030, reveals no PKA hyperphosphorylation of the cardiac ryanodine receptor in canine heart failure. *Circ. Res.* **96**, 847–855 (2005).
40. Fruen, B. R. et al. Regulation of the RYR1 and RYR2 Ca^{2+} release channel isoforms by Ca^{2+}-insensitive mutants of calmodulin. *Biochemistry* **42**, 2740–2747 (2003).
41. Smart, O. S., Neduvelil, J. G., Wang, X., Wallace, B. A. & Sansom, M. S. HOLE: a program for the analysis of the pore dimensions of ion channel structural models. *J. Mol. Graph.* **14**, 354–360 (1996).

METHODS

Expression and purification of GST–FKBP12.6. Because the sequence of porcine FKBP12.6 is not available in the public domain, human FKBP12.6 was applied to pull-down porcine RyR2 (pRyR2)[30]. The complementary DNA of full-length human *FKBP12.6* (also known as *FKBP1B*) was cloned into the pGEX-4T-2 vector with a C-terminal 6×His tag and an N-terminal glutathione *S*-transferase (GST) tag. Protein was overexpressed in the *Escherichia coli* BL21 (DE3) strain at 18 °C for 12–15 h after the addition of 0.2 mM isopropyl-β-D-thiogalactoside (IPTG) to cells with an optical density at 600 nm (OD_{600}) of 1.0. Cells collected by centrifugation were resuspended in lysis buffer (25 mM Tris, pH 8.0, 150 mM NaCl). Cell debris was removed by centrifugation at 22,000*g* for 1 h, and the supernatant was applied to Ni^{2+}-NTA resin (Qiagen). The resin was washed with both W1 buffer (25 mM Tris, pH 8.0, 500 mM NaCl) and W2 buffer (25 mM Tris, pH 8.0, 20 mM imidazole) and eluted with 25 mM Tris, pH 8.0 and 300 mM imidazole. The elution was further purified by anion-exchange chromatography (SOURCE 15Q, GE Healthcare).

Expression and purification of the wild-type CaM and CaM mutant. In mammals, three independent genes (*CALM1*–*CALM3*) with approximately 80% identity[42] are transcribed into at least eight mRNAs that encode identical CaM proteins[43]. It has previously been reported that the first methionine residue of CaM was removed under physiological conditions[44]. The complementary DNA of human *CALM3* without the initial Met was cloned into the pET21 vector with an N-terminal 6×His tag followed by an N-terminal SUMO tag and a stop codon in the C terminus, preventing the translation of a C-terminal 6×His tag in the original pET21 vector (Extended Data Fig. 1a). The expression and purification protocol was similar to that of GST–FKBP12.6 mentioned above. Specifically, the N-terminal 6×His tag and SUMO tag were removed together by the SUMO protease Ulp1p[45] during purification. The CaM protein was further purified by anion-exchange chromatography (SOURCE 15Q, GE Healthcare) using buffer 1 (25 mM Tris, pH 8.0) and buffer 2 (1 M NaCl, 25 mM Tris, pH 8.0). Finally, the protein was applied to size-exclusion chromatography (SEC; Superdex-200, GE Healthcare) in buffer F (20 mM HEPES, pH 7.4, 200 mM NaCl, 0.1% digitonin, 1.3 μg ml^{-1} aprotinin, 1 μg ml^{-1} pepstatin, 5 μg ml^{-1} leupeptin, 0.2 mM PMSF and 2 mM DTT), which is the same as that used for the last-step purification of RyR2. The N-terminal boundary of wild-type CaM was confirmed by N-terminal sequencing (Extended Data Fig. 1b). The expression and purification of the CaM mutant that is deficient in Ca^{2+} binding at all four EF-hand Ca^{2+}-binding sites (E32A, E68A, E105A and E141A) (denoted as CaM-M) were the same as for the wild-type CaM.

Preparation of sarcoplasmic reticulum membranes from porcine heart. The procedures for preparing the membranes of the sarcoplasmic reticulum from porcine hearts were similar to previously described procedures[30]. A single porcine heart was cut into small pieces and then resuspended in five volumes of homogenization buffer A (20 mM HEPES, pH 7.4, 150 mM NaCl, 5 mM EDTA, 1.3 μg ml^{-1} aprotinin, 1 μg ml^{-1} pepstatin, 5 μg ml^{-1} leupeptin and 0.2 mM PMSF). Homogenization was performed in a blender (JYL-C010, Joyoung) for fifteen cycles. The debris was removed by low-speed centrifugation (6,000*g*) for 10 min. The supernatant was further centrifuged at high speed (20,000*g*) for 1 h. The pellet was then resuspended in two volumes of homogenization buffer B (20 mM HEPES, pH 7.4, 1 M NaCl, 1.3 μg ml^{-1} aprotinin, 1 μg ml^{-1} pepstatin, 5 μg ml^{-1} leupeptin, 0.2 mM PMSF and 2 mM DTT) and flash-frozen in liquid nitrogen.

Purification of pRyR2 by GST–FKBP12.6. The pRyR2–FKBP12.6 complex was purified based on previously described procedures[30] with slight modifications. The membrane of the sarcoplasmic reticulum from a single porcine heart was solubilized at 4 °C for 2 h in homogenization buffer B supplemented with 5% CHAPS and 1.25% soy bean lecithin. After solubilization, the final concentration of NaCl in the system was diluted to 200 mM by homogenization buffer B without NaCl. Approximately 5–6 mg of purified GST–FKBP12.6 was then added to the system and further incubated for 1 h at 4 °C. After ultrahigh-speed centrifugation (200,000*g*), the supernatant was loaded onto a GS4B column (GE Healthcare). The resin was washed with buffer similar to the homogenization buffer B, except that the NaCl concentration was 200 mM and 0.1% digitonin was added. The complex was eluted by a solution containing 80 mM Tris, pH 8.0, 200 mM NaCl, 10 mM GSH, 0.1% digitonin, 1.3 μg ml^{-1} aprotinin, 1 μg ml^{-1} pepstatin, 5 μg ml^{-1} leupeptin, 0.2 mM PMSF and 2 mM DTT. The eluted protein was further purified through SEC (Superose 6, 10/300 GL, GE Healthcare) in buffer F. The pRyR2–FKBP12.6 complex fractions were concentrated to approximately 0.1 mg ml^{-1} for electron microscopy sample preparation. Specifically, for the FKBP12.6/apo-CaM sample, 5 mM EDTA, which has no effect on the zinc finger structure of RyR2[30], was included throughout purification of RyR2. For the CHAPS- and DOPC-treated FKBP12.6/ATP/caffeine/low-[Ca^{2+}] and CHAPS- and DOPC-treated FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/Ca^{2+}-CaM samples, the proteins were extracted only by CHAPS and washed and eluted by a buffer containing 0.5% CHAPS plus 0.002% DOPC. The eluted proteins were further purified through SEC in buffer F except 0.1% digitonin was replaced by 0.25% CHAPS plus 0.001% DOPC. For the PCB95/low-[Ca^{2+}]/Ca^{2+}-CaM sample, the proteins were purified using GST–FKBP12 as a bait and the RyR2–FKBP12 (containing GST–FKBP12) complex fell apart during SEC purification[30].

Cryo-EM sample preparation. The cryo-EM samples of RyR2–CaM complexes were prepared as follows. FKBP12.6/apo-CaM: 5 mM EDTA was added to CaM (in buffer F) before sample preparation, and CaM with a final concentration of 250 μM was added to RyR2 (in buffer F plus 5 mM EDTA). FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/CaM-M: CaM-M (in buffer F) and RyR2 (in buffer F) were separately added with 20 μM Ca^{2+} (low-Ca^{2+} concentration), 5 mM ATP and 5 mM caffeine and CaM-M with a final concentration of 250 μM was added to RyR2. FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/Ca^{2+}-CaM: CaM (in buffer F) and RyR2 (in buffer F) were separately added to 20 μM Ca^{2+}, 5 mM ATP and 5 mM caffeine and CaM with a final concentration of 2.5 μM was added to RyR2. The other samples were prepared by the same procedure. Note that 5 mM Ca^{2+} represents high-Ca^{2+} concentration. Vitrobot Mark IV (FEI) was used for the preparation of cryo-EM grids. The procedures for preparing the eight samples were the same. Aliquots (3 μl each) of pRyR2 samples were placed on glow-discharged lacey carbon grids (Ted Pella). Grids were blotted for 2 s and flash-frozen in liquid ethane. Owing to the presence of high concentrations of Ca^{2+} in the filter paper used for blotting, it has previously been reported that the final concentration of free Ca^{2+} may be much higher than those used during sample preparation[46]. The low- and high-Ca^{2+} concentrations presented here only indicate those used during sample preparation and may be lower than the true Ca^{2+} concentrations.

Cryo-EM image acquisition. With regard to the FKBP12.6/ATP/caffeine/low-[Ca^{2+}], FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/CaM-M, CHAPS- and DOPC-treated FKBP12.6/ATP/caffeine/low-[Ca^{2+}], CHAPS- and DOPC-treated FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/Ca^{2+}-CaM, FKBP12.6/ATP/caffeine/high-[Ca^{2+}]/Ca^{2+}-CaM and PCB95/low-[Ca^{2+}]/Ca^{2+}-CaM datasets, grids were transferred to a Titan Krios (Thermo Fisher Scientific) electron microscope operating at 300 kV equipped with a Cs-corrector (Thermo Fisher Scientific), Gatan K2 Summit detector and GIF Quantum energy filter. Zero-loss movie stacks were automatically collected using AutoEMationII[47,48] with a slit width of 20 eV on the energy filter and a defocus range from −1.3 μm to −1.7 μm in super-resolution mode at a nominal magnification of 105,000×. Each stack was exposed for 5.6 s with an exposure time of 0.175 s per frame, resulting in 32 frames per stack. The total dose was approximately 50 e^{-} $Å^{-2}$ for each stack. The stacks were motion-corrected with MotionCor2[49] and binned twofold, resulting in a pixel size of 1.091 Å per pixel. With regard to the FKBP12.6/apo-CaM and FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/Ca^{2+}-CaM datasets, micrographs were collected using a Gatan K2 Summit detector mounted on a Titan Krios electron microscope (FEI Company) operating at 300 kV and equipped with a GIF Quantum energy filter (slit width 20 eV). Micrographs were recorded in the super-resolution mode with a normal magnification of 105,000 ×, resulting in a calibrated pixel size of 0.669 Å. Each stack of 32 frames was exposed for 8 s, with an exposing time of 0.25 s per frame. The total dose rate was about 45.6 e^{-} $Å^{-2}$ for each stack. All 32 frames in each stack were motion-corrected with MotionCor2 and binned to a pixel size of 1.338 Å. The defocus value of each image was set from −0.8 μm to −1.8 μm. In addition, dose weighting was performed[50]. The defocus values were estimated with Gctf[51].

Image processing. Image-processing procedures were similar to those previously reported[30]. Diagrams of the procedures used in data processing are presented in Extended Data Fig. 2. For the FKBP12.6/apo-CaM dataset, 1,180,104 particles were picked from 7,800 micrographs by RELION 2.0[52] using templates low-pass-filtered to 20 Å to limit reference bias. After two rounds of two-dimensional classification, 832,833 particles were selected and subjected to global angular search three-dimensional classification using RELION 2.0 with one class and a step size of 7.5°. The electron microscopy map of the previously published open structure of RyR2[30], which was low-pass-filtered to 60 Å, was used as the initial model. After global angular search three-dimensional classification, the particles were further subjected to three-dimensional classification with 10 classes and a local angular search step of 3.75°. The local angular search three-dimensional classification was performed several times with the output from different iterations of the global angular search three-dimensional classification as input. After the merging of all good classes and removal of the duplicated particles, the particles were subjected to three-dimensional autorefinement using THUNDER software[53]. The final particle number for the three-dimensional autorefinement was 208,715, resulting in a 3.6 Å resolution map after post-processing. The same procedures were performed for the other datasets. The resolution was estimated with the gold-standard Fourier shell correlation 0.143 criterion[54] with the high-resolution noise-substitution method[55].

Model building and structure refinement. The model of the RyR2 open structure (PDB code 5GOA)[30] was fitted into the maps of the eight conditions by Chimera[56] and manually adjusted in COOT[57]. FKBP12 from the rabbit RyR1/FKBP12 complex structure (PDB code 3J8H)[58] was used for homologous model building of FKBP12.6. The apo-CaM from the crystal structure 3WFN was fitted into the maps obtained in the presence of CaM-M or apo-CaM and manually adjusted in

COOT. Similarly, the crystal structure of Ca^{2+}-CaM in complex with the RyR1 peptide (PDB code 2BCX) was fitted into the maps obtained in the presence of Ca^{2+}-CaM and manually adjusted in COOT. Structure refinement was performed using PHENIX[59] in real space with restrained secondary structure and geometry. The statistics of the three-dimensional reconstruction and model refinement are summarized in Extended Data Table 1.

Evaluation of the conformations of N- and C-lobes of CaM. Different conformations of N- and C-lobes were docked into the electron microscopy reconstruction for FKBP12.6/ATP/caffeine/high-[Ca^{2+}]/Ca^{2+}-CaM using the 'Fit in Map' tool of Chimera, selecting the options that include 'Real-time correlation', '7-Å resolution of Use map simulated from atoms', 'Use only data above contour level from first map', 'Optimize correlation', 'Correlation calculated about mean data value', 'Allow rotation and shift' and 'Move whole molecules'.

Site-directed mutagenesis. Point mutations in mouse *Ryr2* and in human *CALM1* were generated with the overlap extension method using PCR. In brief, a EcoRV/HpaI DNA fragment containing the RyR2(Y2156A) mutation and an AgeI/SalI fragment containing the RyR2(V3599A), RyR2(W3587A) or RyR2(L3590A) mutations were obtained by overlapping PCR and used to replace the corresponding wild-type fragment in the NheI/BsiWI fragment of *Ryr2*, The mutated NheI/BsiWI fragment was then used to replace the corresponding wild-type fragment in the full-length *Ryr2* cDNA in pcDNA5. A HindIII/XhoI full-length *CALM1* DNA fragment containing various point mutations was generated by overlapping PCR, which was then subcloned into pcDNA3. All point mutations in *Ryr2* and *CALM1* were confirmed by DNA sequencing.

Generation of stable, inducible cell lines expressing RyR2(WT) and mutants. Stable, inducible HEK293 cell lines expressing RyR2(WT), RyR2(Y2156A), RyR2(V3599A), RyR2(W3587A) and RyR2(L3590A) were generated using the Flp-In T-REx Core Kit from Invitrogen. These cell lines were not authenticated. These cells tested negative for mycoplasma contamination. In brief, Flp-In T-REx HEK293 cells were co-transfected with the inducible expression vector pcDNA5/FRT/TO containing the $Ryr2^{WT}$ or *Ryr2*-mutant cDNA and the pOG44 vector encoding the Flp recombinase in 1:5 ratios using the calcium phosphate precipitation method. The transfected cells were washed with phosphate buffered saline (PBS; 137 mM NaCl, 8 mM Na_2HPO_4, 1.5 mM KH_2PO_4 and 2.7 mM KCl, pH 7.4) 24 h after transfection followed by a change into fresh medium for 24 h. The cells were then washed again with PBS, collected and plated onto new dishes. After the cells had attached (around 4 h), the growth medium was replaced with a selection medium containing 200 μg ml^{-1} hygromycin (Invitrogen). The selection medium was changed every 3–4 days until the desired number of cells was grown. The hygromycin-resistant cells were pooled, aliquoted (1 ml) and stored at −80 °C. These positive cells are believed to be isogenic, because the integration of *Ryr2* cDNA is mediated by the Flp recombinase at a single FRT site.

Single-cell luminal Ca^{2+} imaging. Luminal Ca^{2+} levels in $Ryr2^{WT}$- or *Ryr2*-mutant-expressing HEK293 cells transfected with or without CaM(WT) or CaM mutants were measured using single-cell Ca^{2+} imaging and the fluorescence resonance energy transfer (FRET)-based endoplasmic-reticulum luminal Ca^{2+}-sensitive chameleon protein D1ER as previously described[60,61]. The cells were grown to 95% confluence in a 75-cm^2 flask, dissociated with PBS (137 mM NaCl, 8 mM Na_2HPO_4, 1.5 mM KH_2PO_4 and 2.7 mM KCl, pH 7.4) and plated on glass coverslips placed on tissue culture dishes at approximately 10% confluence 18–20 h before transfection with cDNA for D1ER and cDNAs for CaM(WT) or CaM mutants using the calcium phosphate precipitation method. After transfection for 24 h, the growth medium was then changed to an induction medium containing 1 μg ml^{-1} tetracycline. After induction for around 22 h, the coverslip was mounted onto an inverted microscope (Nikon TE2000-S) and the cells on the coverslip were perfused continuously with Krebs–Ringer–HEPES buffer (125 mM NaCl, 5 mM KCl, 1.2 mM KH_2PO_4, 6 mM glucose, 1.2 mM $MgCl_2$ and 25 mM HEPES, pH 7.4) containing various concentrations of $CaCl_2$ (0, 1 and 2 mM) to induce SOICR, followed by the addition of 1.0 mM tetracaine, which was used to estimate the store capacity, and caffeine (20 mM), which was used to estimate the minimum store level by depleting the endoplasmic-reticulum Ca^{2+} stores at room temperature (23 °C). Images were captured with Compix Simple PCI 6 software every 2 s using the Nikon TE2000-S inverted microscope equipped with an S-Fluor 20×/0.75 NA objective. The filters used for D1ER imaging were $\lambda_{ex} = 436 \pm 20$ nm for CFP and $\lambda_{ex} = 500 \pm 20$ nm for YFP, and $\lambda_{em} = 465 \pm 30$ nm for CFP and $\lambda_{em} = 535 \pm 30$ nm for YFP with a dichroic mirror (500 nm). The amount of FRET in individual cells was determined from the ratio of the light emission at 535 and 465 nm. F_{SOICR} is defined as the FRET level at which SOICR occurs, and F_{termi} is defined as the FRET level at which SOICR terminates. The maximum FRET signal F_{max} is defined as the FRET level after tetracaine treatment. The minimum FRET signal F_{min} is defined as the FRET level after caffeine treatment. The termination and activation thresholds of SOICR in individual cells were determined using the equations shown in Extended Data Fig. 7a. The store capacity is calculated by subtracting F_{min} from F_{max}. Individual data points represent the average measurements of around 10–30 cells from one coverslip in one set of experiment. The number of experiments and coverslips for each condition is used as the sample size for data analyses.

Single-channel recordings in planar lipid bilayers. Recombinant RyR2(WT) and mutant channels were purified from cell lysates prepared from HEK293 cells transfected with the $Ryr2^{WT}$ or *Ryr2*-mutant (Y2156A, V3599A, W3587A or L3590A) cDNA by sucrose density gradient centrifugation as previously described[62]. Heart phosphatidylethanolamine (50%) and brain phosphatidylserine (50%) (Avanti Polar Lipids), dissolved in chloroform, were combined and dried under nitrogen gas and resuspended in 30 μl of *n*-decane at a concentration of 12 mg lipid per ml. Bilayers were formed across a 250-μm hole in a Delrin partition separating two chambers. The *trans* chamber (800 μl) was connected to the head stage input of an Axopatch 200A amplifier (Axon Instruments). The *cis* chamber (1.2 ml) was held at virtual ground. A symmetrical solution containing 250 mM KCl and 25 mM HEPES, pH 7.4 was used for all recordings, unless indicated otherwise. A 4-μl aliquot (around 1 μg protein) of the sucrose density gradient-purified recombinant RyR2(WT) or mutant channels was added to the *cis* chamber. Spontaneous channel activity was always tested for sensitivity to EGTA and Ca^{2+}. The chamber to which the addition of EGTA inhibited the activity of the incorporated channel presumably corresponds to the cytosolic side of the Ca^{2+} release channel. The direction of single channel currents was always measured from the luminal to the cytosolic side of the channel, unless mentioned otherwise. Recordings were filtered at 2,500 Hz. Data analyses were carried out using the pCLAMP 8.1 software package (Axon Instruments). Free Ca^{2+} concentrations were calculated using a computer program that has previously been described[63].

Statistical analysis. Data are mean ± s.e.m., derived from independent samples or independent experiments. All experiments were performed with at least five biological replicates. The GraphPad Prism 8.1 software was used to test for differences between groups. We used Student's *t*-test (paired, two-tailed) or one-way ANOVA with a Dunnett's post hoc test. $P < 0.05$ was considered to be statistically significant.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Atomic coordinates and electron microscopy density maps of the following structures have been deposited in the PDB (http://www.rcsb.org) and the Electron Microscopy Data Bank (EMDB https://www.ebi.ac.uk/pdbe/emdb/). FKBP12.6/apo-CaM (PDB, 6JI8; EMDB, EMD-9833), FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/CaM-M (PDB, 6JII; EMDB, EMD-9834), FKBP12.6/ATP/caffeine/low-[Ca^{2+}] (PDB, 6JI0; EMDB, EMD-9831), FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/Ca^{2+}-CaM (PDB, 6JIU; EMDB, EMD-9836), CHAPS- and DOPC-treated FKBP12.6/ATP/caffeine/low-[Ca^{2+}] (PDB, 6JRR; EMDB, EMD-9879), CHAPS- and DOPC-treated FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/Ca^{2+}-CaM (PDB, 6JRS; EMDB, EMD-9880), FKBP12.6/ATP/caffeine/high-[Ca^{2+}]/Ca^{2+}-CaM (PDB, 6JIY; EMDB, EMD-9837) and PCB95/low-[Ca^{2+}]/Ca^{2+}-CaM (PDB, 6JV2; EMDB: EMD-9889) complexes. Source Data for Fig. 2e, f and Extended Data Figs. 1c, 6f, 7d–h are available in the online version of the paper. All other data are available from the corresponding authors upon reasonable request.

42. Fischer, R. et al. Multiple divergent mRNAs code for a single human calmodulin. *J. Biol. Chem.* **263**, 17055–17062 (1988).
43. Kortvely, E. & Gulya, K. Calmodulin, and various ways to regulate its activity. *Life Sci.* **74**, 1065–1070 (2004).
44. Sasagawa, T. et al. Complete amino acid sequence of human brain calmodulin. *Biochemistry* **21**, 2565–2569 (1982).
45. Hirano, H., Kobayashi, J. & Matsuura, Y. Structures of the karyopherins Kap121p and Kap60p bound to the nuclear pore-targeting domain of the SUMO protease Ulp1p. *J. Mol. Biol.* **429**, 249–260 (2017).
46. Paknejad, N. & Hite, R. K. Structural basis for the regulation of inositol trisphosphate receptors by Ca^{2+} and IP_3. *Nat. Struct. Mol. Biol.* **25**, 660–668 (2018).
47. Fan, X. et al. Near-atomic resolution structure determination in over-focus with volta phase plate by Cs-corrected cryo-EM. *Structure* **25**, 1623–1630 (2017).
48. Lei, J. & Frank, J. Automated acquisition of cryo-electron micrographs for single particle reconstruction on an FEI Tecnai electron microscope. *J. Struct. Biol.* **150**, 69–80 (2005).
49. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
50. Grant, T. & Grigorieff, N. Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6. *eLife* **4**, e06980 (2015).
51. Zhang, K. Gctf: real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
52. Kimanius, D., Forsberg, B. O., Scheres, S. H. & Lindahl, E. Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *eLife* **5**, e18722 (2016).
53. Hu, M. et al. A particle-filter framework for robust cryo-EM 3D reconstruction. *Nat. Methods* **15**, 1083–1089 (2018).

54. Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol*. **333**, 721–745 (2003).
55. Chen, S. et al. High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. *Ultramicroscopy* **135**, 24–35 (2013).
56. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem*. **25**, 1605–1612 (2004).
57. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
58. Yan, Z. et al. Structure of the rabbit ryanodine receptor RyR1 at near-atomic resolution. *Nature* **517**, 50–55 (2015).
59. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
60. Palmer, A. E., Jin, C., Reed, J. C. & Tsien, R. Y. Bcl-2-mediated alterations in endoplasmic reticulum Ca^{2+} analyzed with an improved genetically encoded fluorescent sensor. *Proc. Natl Acad. Sci. USA* **101**, 17404–17409 (2004).
61. Jones, P. P. et al. Endoplasmic reticulum Ca^{2+} measurements reveal that the cardiac ryanodine receptor mutations linked to cardiac arrhythmia and sudden death alter the threshold for store-overload-induced Ca^{2+} release. *Biochem. J.* **412**, 171–178 (2008).
62. Jiang, D. et al. Enhanced store overload-induced Ca^{2+} release and channel sensitivity to luminal Ca^{2+} activation are common defects of RyR2 mutations linked to ventricular tachycardia and sudden death. *Circ. Res*. **97**, 1173–1181 (2005).
63. Fabiato, A. & Fabiato, F. Calculator programs for computing the composition of the solutions containing multiple metals and ligands used for experiments in skinned muscle cells. *J. Physiol*. **75**, 463–505 (1979).

Acknowledgements We thank X. Li for technical support for electron microscopy image acquisition; the Tsinghua University Branch of China National Center for Protein Sciences (Beijing) for providing the cryo-EM facility support; the computational facility support on the cluster of Bio-Computing Platform (Tsinghua University Branch of China National Center for Protein Sciences Beijing) and the 'Explorer 100' cluster system of Tsinghua National Laboratory for Information Science and Technology; M. T. Overgaard for providing the CaM protein for single channel studies. This work was funded by the National Key R&D Program (2016YFA0500402) and the National Key Basic Research (973) Program (2015CB910101) from Ministry of Science and Technology of China and the National Natural Science Foundation of China (projects 31621092, 31630017 and 81861138009). N.Y. is supported by the Shirley M. Tilghman endowed professorship from Princeton University. This work was also supported by research grants from the Heart and Stroke Foundation of Canada, the Canadian Institutes of Health Research and the Heart and Stroke Foundation Chair in Cardiovascular Research (S.R.W.C.).

Author contributions D.G. and N.Y. conceived the project. D.G., X.C. and G.Z. prepared the electron microscopy samples. D.G., X.C., G.H. and J.L. conducted the cryo-EM analysis. J.W., L.Z. and R.W. performed the functional experiments. All authors contributed to data analysis. D.G., S.R.W.C. and N.Y. wrote the manuscript.

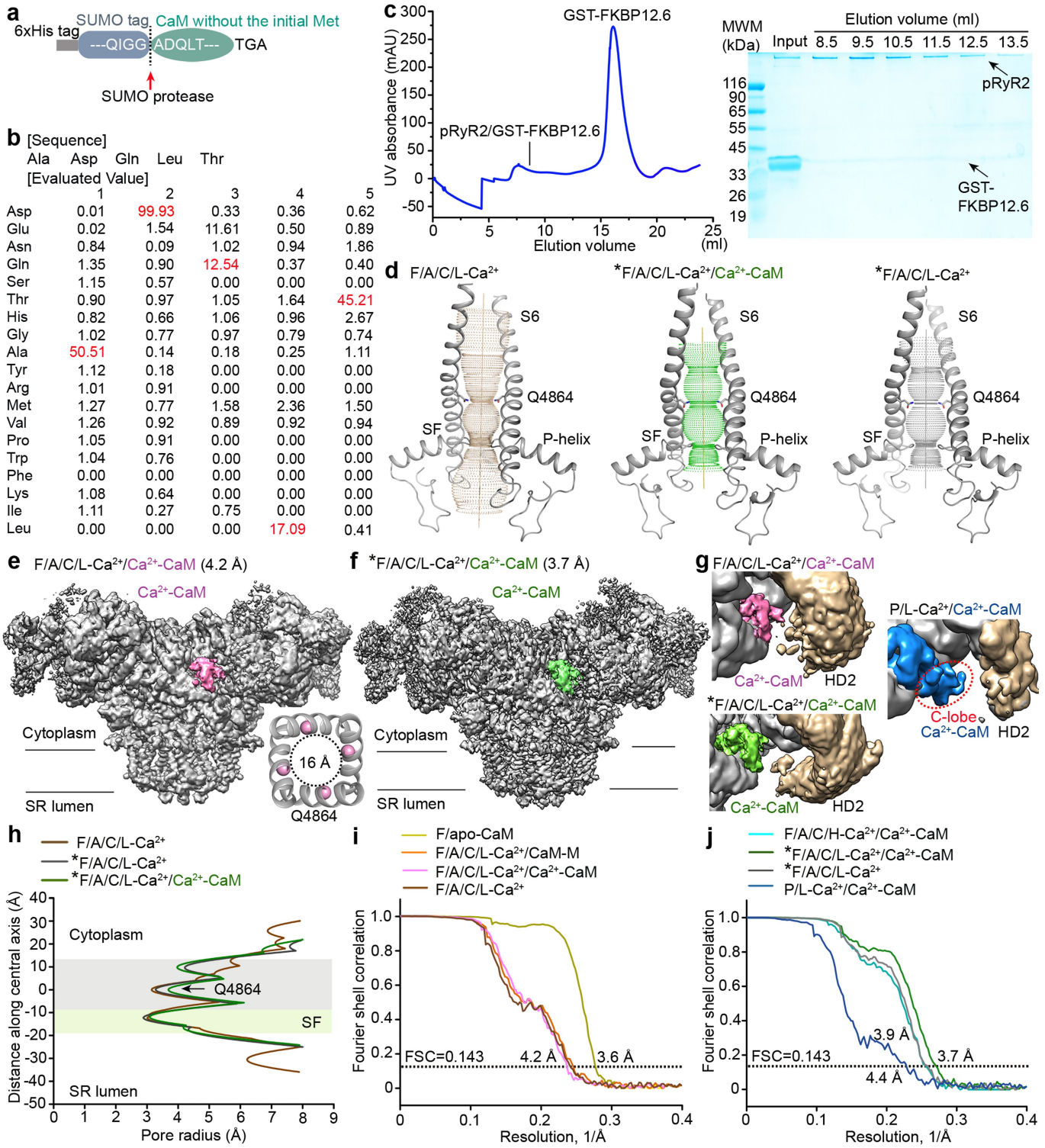
Competing interests The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1377-y.

Correspondence and requests for materials should be addressed to D.G., S.R.W.C. or N.Y.

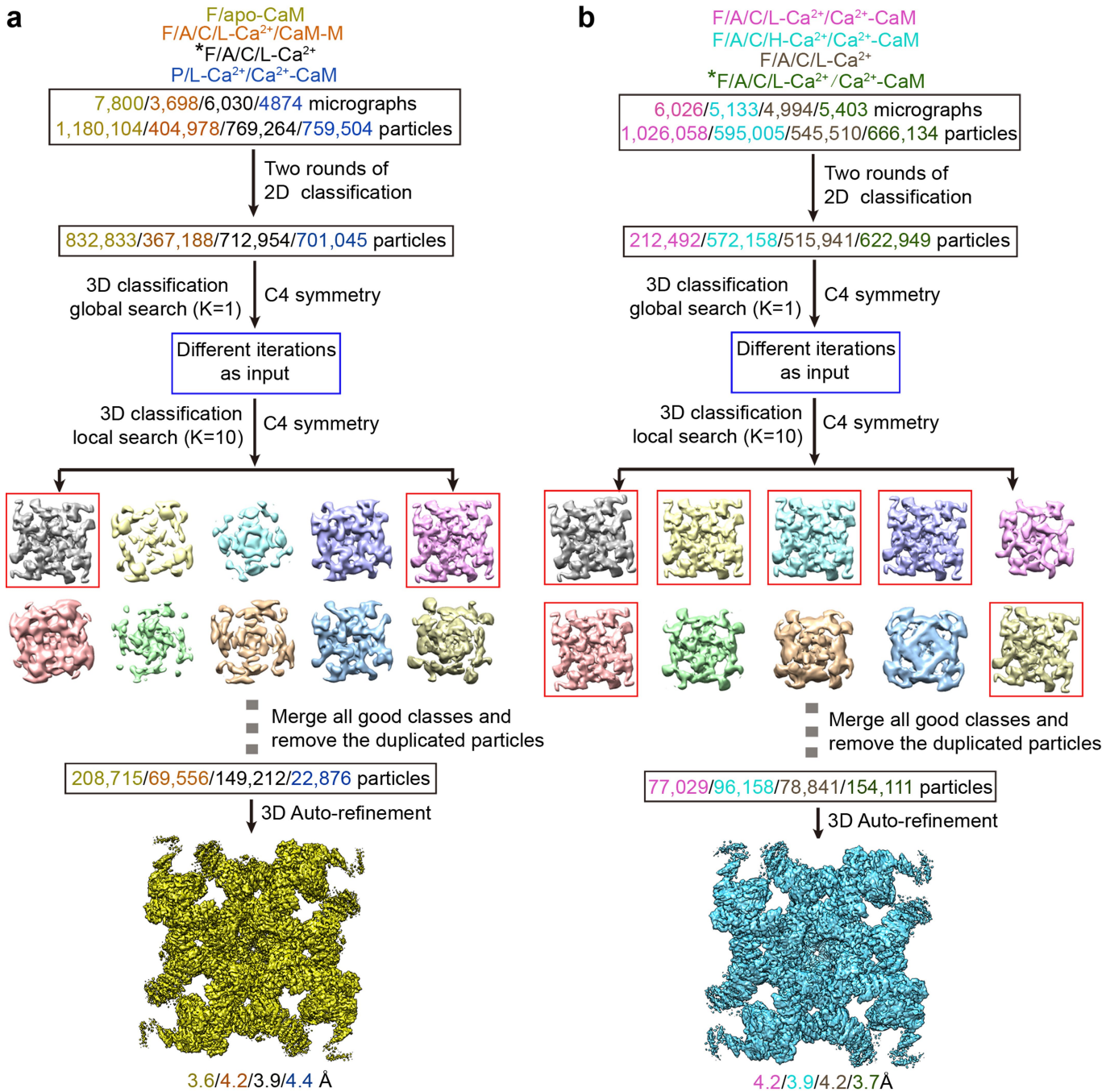
Peer review information *Nature* thanks Youxing Jiang, Filip Van Petegem and the other anonymous reviewer(s) for their contribution to the peer review of this work.



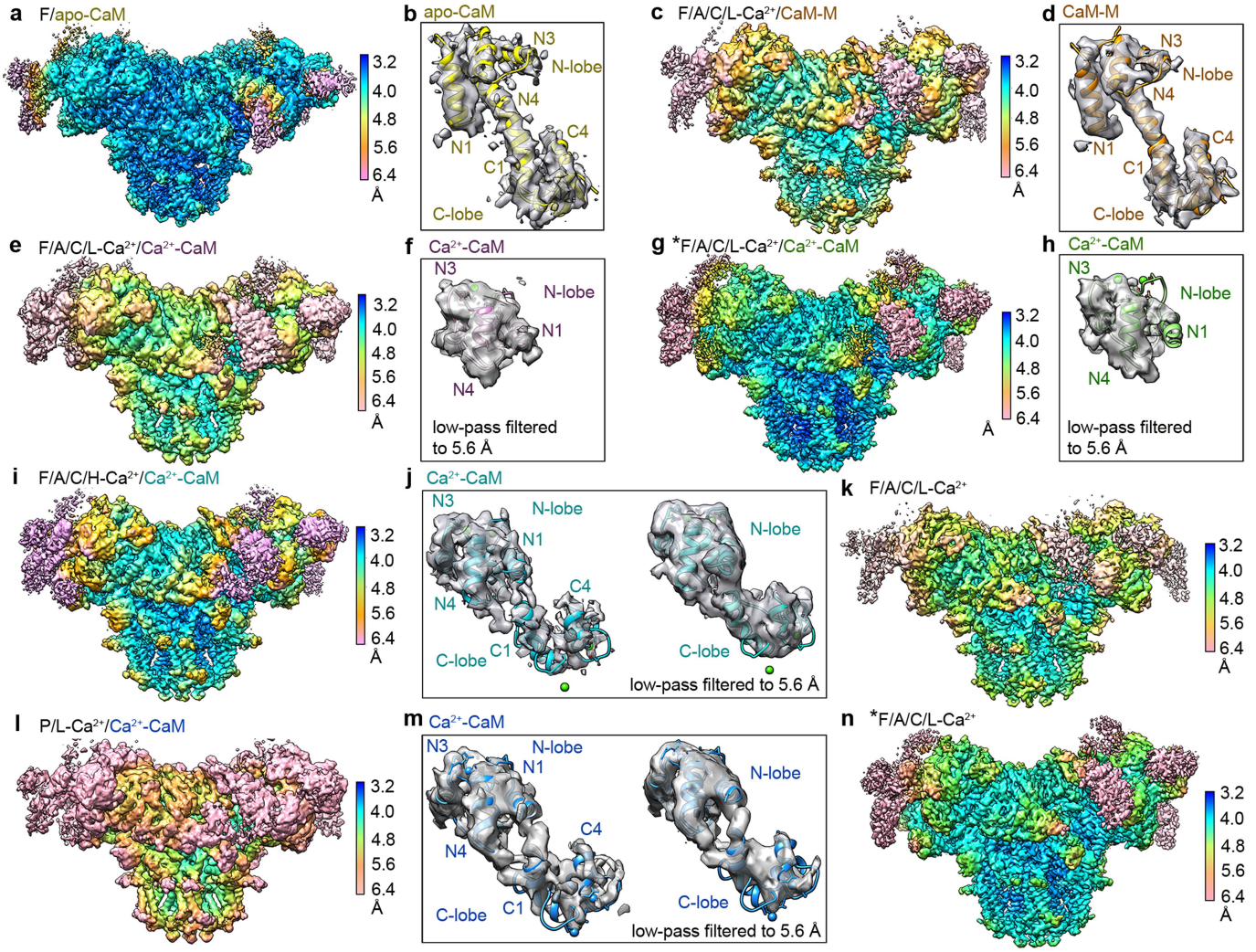
	1	2	3	4	5
Asp	0.01	99.93	0.33	0.36	0.62
Glu	0.02	1.54	11.61	0.50	0.89
Asn	0.84	0.09	1.02	0.94	1.86
Gln	1.35	0.90	12.54	0.37	0.40
Ser	1.15	0.57	0.00	0.00	0.00
Thr	0.90	0.97	1.05	1.64	45.21
His	0.82	0.66	1.06	0.96	2.67
Gly	1.02	0.77	0.97	0.79	0.74
Ala	50.51	0.14	0.18	0.25	1.11
Tyr	1.12	0.18	0.00	0.00	0.00
Arg	1.01	0.91	0.00	0.00	0.00
Met	1.27	0.77	1.58	2.36	1.50
Val	1.26	0.92	0.89	0.92	0.94
Pro	1.05	0.91	0.00	0.00	0.00
Trp	1.04	0.76	0.00	0.00	0.00
Phe	0.00	0.00	0.00	0.00	0.00
Lys	1.08	0.64	0.00	0.00	0.00
Ile	1.11	0.27	0.75	0.00	0.00
Leu	0.00	0.00	0.00	17.09	0.41

Extended Data Fig. 1 | See next page for caption.

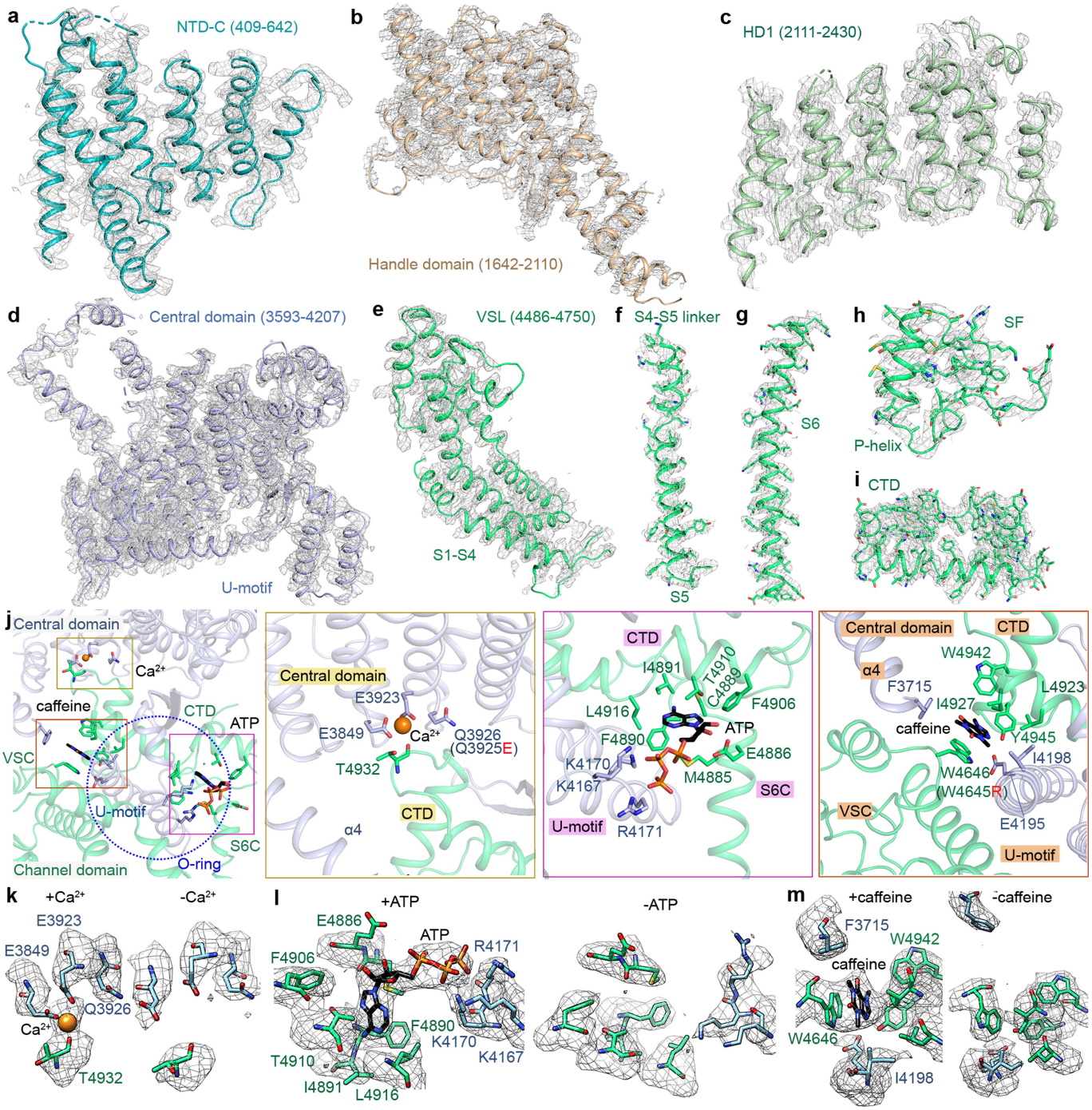
Extended Data Fig. 1 | Protein purification and structural determination. a, Schematic of vector construction for recombinant expression of CaM without the N-terminal Met. **b**, N-terminal sequencing confirmed removal of the initial Met. **c**, SEC purification of the affinity-purified complex of pRyR2–FKBP12.6 (containing GST–FKBP12.6). The experiment was repeated five times independently with similar results. Peak fractions were resolved by SDS–PAGE and visualized by Coomassie blue staining (Supplementary Fig. 1). UV, ultraviolet. MWM, molecular weight marker. **d**, The channel is open in the presence of ATP, caffeine and Ca^{2+} under both digitonin and CHAPS-and-DOPC (indicated by an asterisk) conditions. The pore of CHAPS- and DOPC-treated FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/Ca^{2+}-CaM remains open after Ca^{2+}-CaM loading. The ion-conduction path, calculated by HOLE, is illustrated by dots in each structure. A, ATP; C, caffeine; Ca^{2+}-CaM, Ca^{2+}-bound CaM; CaM-M, a Ca^{2+}-binding-deficient CaM mutant that mimics apo-CaM; F, FKBP12.6; L-Ca^{2+}, low Ca^{2+} concentration. **e**, Overall electron microscopy map of the FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/Ca^{2+}-CaM complex. Inset, the cytoplasmic view of the channel gate. Because the side chains of the Gln4864-gating residues are not well-resolved, the distance between the C_α atoms of Gln4864 gating in the diagonal protomers is shown in the dashed circle. The density corresponding to CaM was generated from the map that was low-pass-filtered to 5.6 Å with a contour level of 0.015; the other regions were from the 4.2 Å map with a contour level of 0.023. **f**, Overall electron microscopy map of the CHAPS- and DOPC-treated FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/Ca^{2+}-CaM complex. The density corresponding to CaM was generated from the map that was low-pass-filtered to 5.6 Å with a contour level of 0.013; the other regions were from the 3.7 Å map with a contour level of 0.021. **g**, Although the concentrations of Ca^{2+}-CaM are the same in these three conditions, only the N-lobe can be resolved in the FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/Ca^{2+}-CaM and CHAPS- and DOPC-treated FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/Ca^{2+}-CaM RyR2 structures. The reason may be that the HD2 in these two structures presents a steric hindrance for C-lobe binding. P, PCB95. **h**, The corresponding pore radii of the three structures are plotted. **i, j**, Gold-standard Fourier shell correlation curves for electron microscopy maps of the eight datasets. H-Ca^{2+}, high Ca^{2+} concentration.



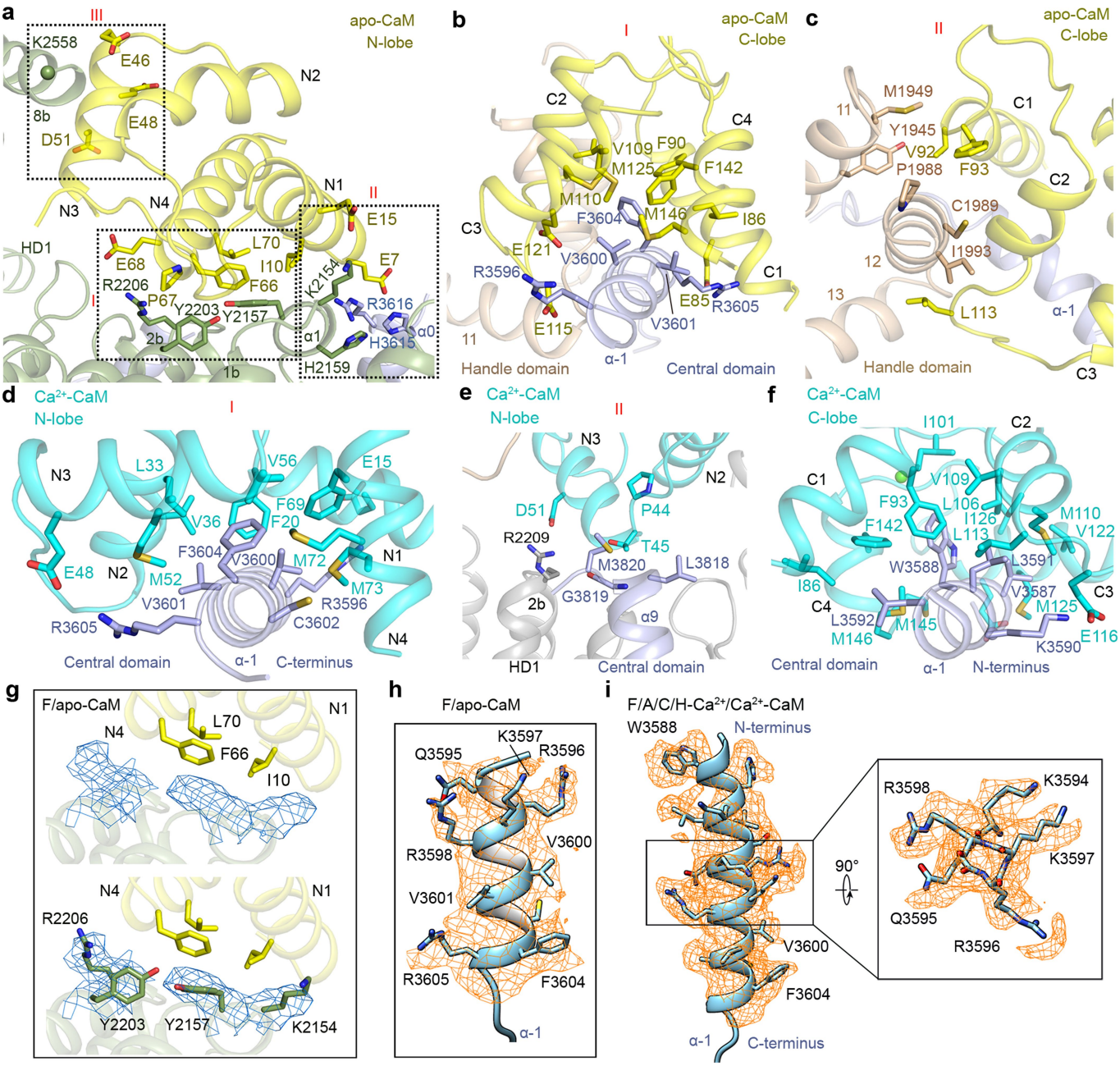
Extended Data Fig. 2 | Flow chart for cryo-EM data processing. See Methods for details. **a**, Data processing of the FKBP12.6/apoCaM, FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/CaM-M, CHAPS- and DOPC-treated FKBP12.6/ATP/caffeine/low-[Ca^{2+}], and PCB95/low-[Ca^{2+}]/Ca^{2+}-CaM datasets. **b**, Data processing of the FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/Ca^{2+}-CaM, FKBP12.6/ATP/caffeine/high-[Ca^{2+}]/Ca^{2+}-CaM, FKBP12.6/ATP/caffeine/low-[Ca^{2+}], and CHAPS- and DOPC-treated FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/Ca^{2+}-CaM datasets.



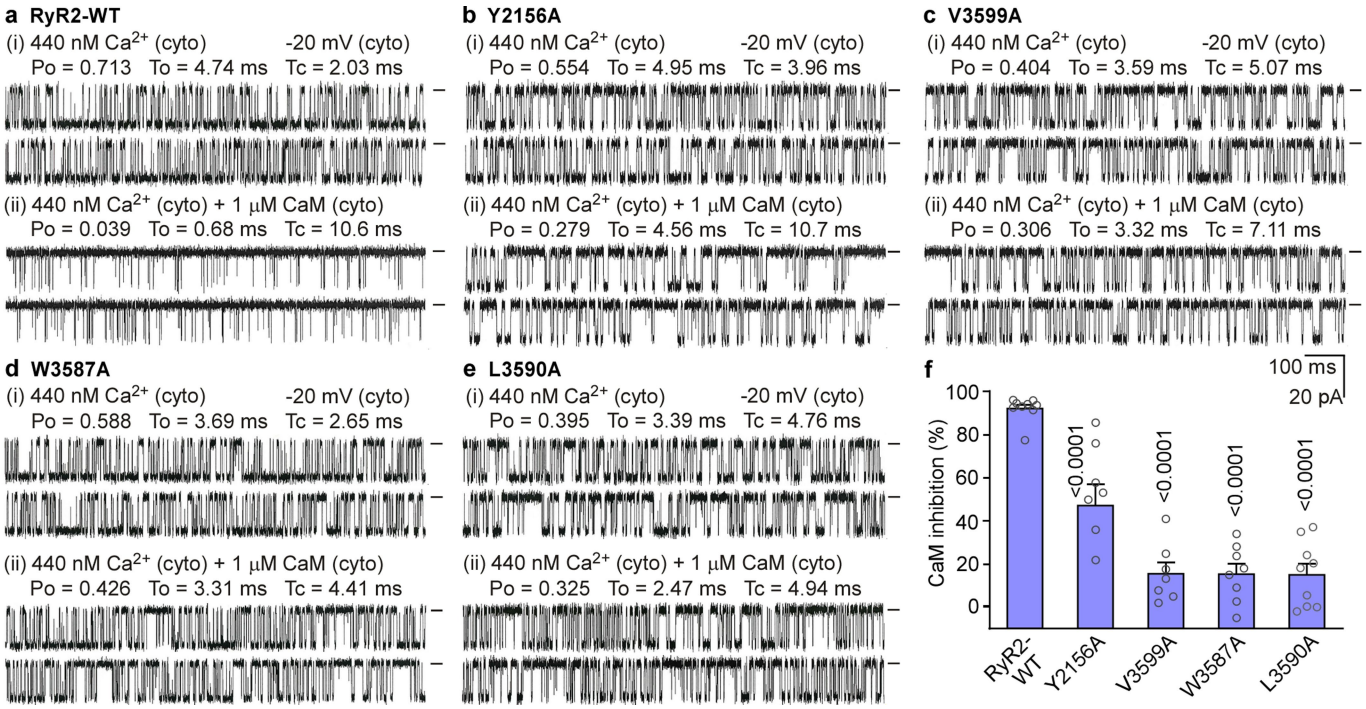
Extended Data Fig. 3 | Local resolution maps of the eight reconstructions. a, **c**, **e**, **g**, **i**, **k**, **l**, **n**, The local resolution maps estimated with RELION 2.0. All electron microscopy maps were generated in Chimera and contoured at levels of 0.027 (**a**), 0.022 (**c**), 0.023 (**e**), 0.021 (**g**), 0.02 (**i**), 0.021 (**k**), 0.015 (**l**) and 0.021 (**n**). **b**, Electron microscopy map of apo-CaM from the reconstruction shown in **a**. **d**, Electron microscopy map of CaM-M. **f**, **h**, The electron microscopy densities of Ca^{2+}-CaM were generated from the maps of FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/Ca^{2+}-CaM and CHAPS- and DOPC-treated FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/Ca^{2+}-CaM that were low-pass-filtered to 5.6 Å with contour levels of 0.015 (**f**) and 0.013 (**h**), respectively. **j**, **m**, Electron microscopy densities of Ca^{2+}-CaM in FKBP12.6/ATP/caffeine/high-[Ca^{2+}]/Ca^{2+}-CaM (**j**) and PCB95/low-[Ca^{2+}]/Ca^{2+}-CaM (**m**). The densities of both lobes were resolved in the map, although the N-lobe was resolved better than the C-lobe.



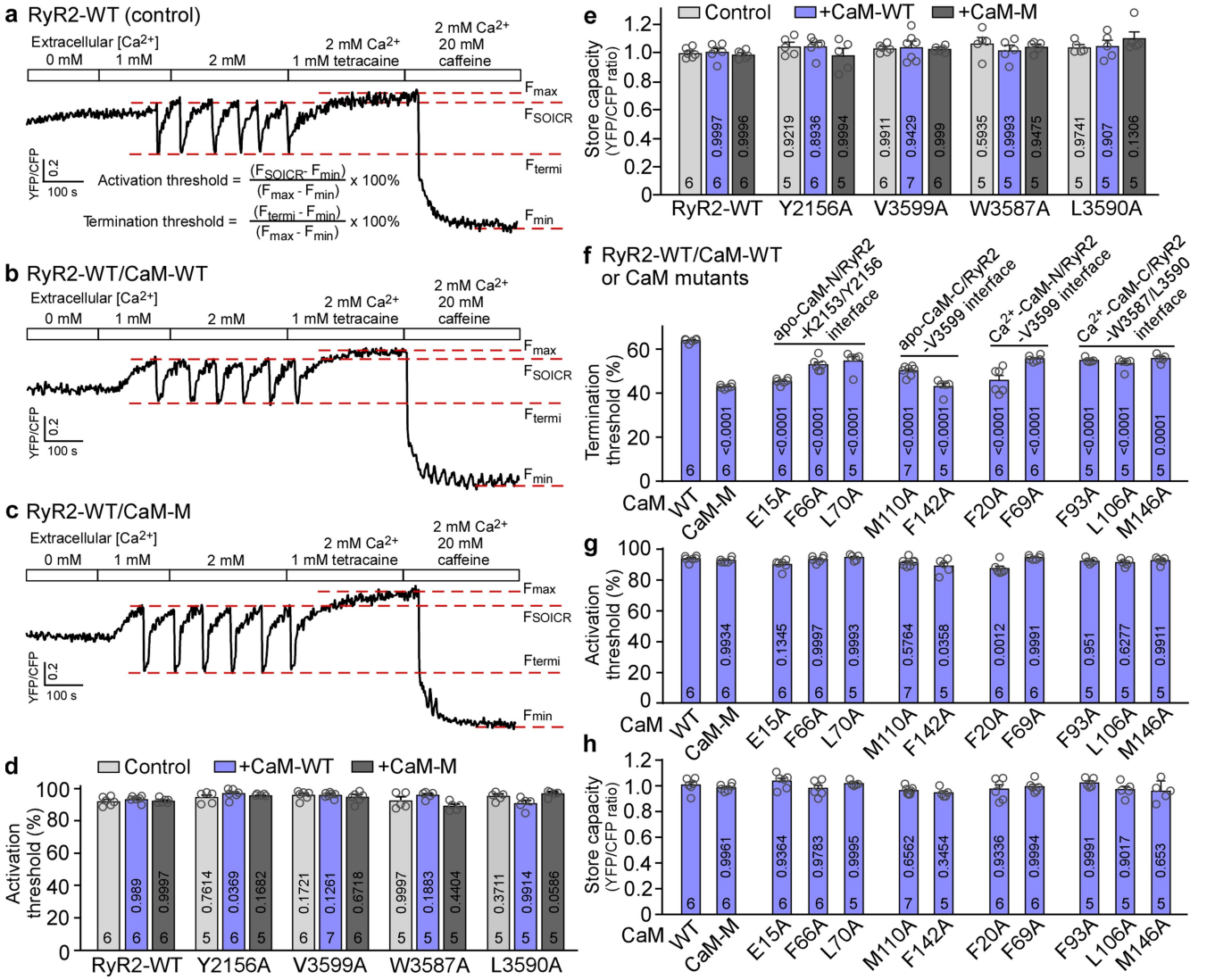
Extended Data Fig. 4 | Representative local electron microscopy maps for FKBP12.6/apo-CaM and densities of the binding sites for Ca^{2+}, ATP and caffeine in FKBP12.6/ATP/caffeine/high-[Ca^{2+}]/Ca^{2+}-CaM. a–i, The electron microscopy maps for the representative segments of RyR2. All of the maps were contoured at 5.5σ. **j,** The binding sites for Ca^{2+}, ATP and caffeine in RyR2. The blue, dotted circle indicates the O-ring that is formed by the C-terminal subdomain (CTD), cytoplasmic subdomain in the voltage-sensor-like domain (VSC) and the cytoplasmic portion of S6 (S6C). Ca^{2+} is located in the cleft that is formed by the central domain and C-terminal subdomain. ATP is located in a pocket formed by the U-motif, C-terminal subdomain and S6C. Caffeine is located at the interface formed by the U-motif, helix α4, C-terminal subdomain and voltage-sensor-like domain. The red letter indicates a disease-causing variant. **k–m,** The local densities of the Ca^{2+}-, ATP- and caffeine-binding sites. The electron microscopy maps of the FKBP12.6/ATP/caffeine/high-[Ca^{2+}]/Ca^{2+}-CaM (+ATP, caffeine and Ca^{2+}) and FKBP12.6/apo-CaM (−ATP, caffeine and Ca^{2+}) RyR2 structures are shown. All of the electron microscopy maps were contoured at a level of 0.029.



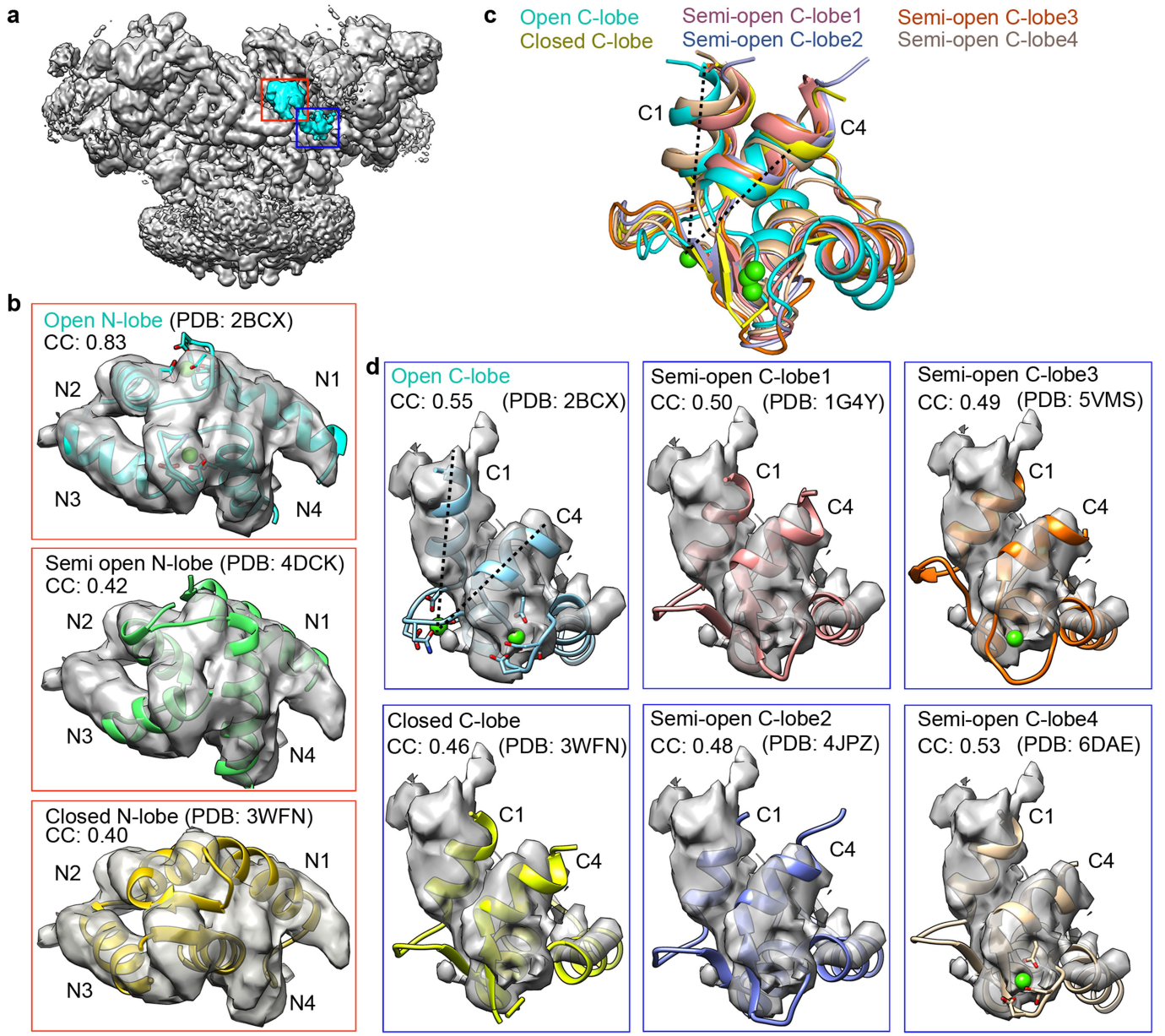
Extended Data Fig. 5 | Binding interfaces between CaM and RyR2. **a**, Three interfaces are formed between the N-lobe of apo-CaM and RyR2. HD1 serves as the major binding site for the N-lobe. The C_α atom of Lys2558 is shown as a sphere. **b**, **c**, Two interfaces are formed between the C-lobe of apo-CaM and RyR2. Helix α−1 is the major binding site of the C-lobe. **d**, **e**, The interfaces between the N-lobe of Ca^{2+}-CaM and RyR2. **f**, The interface between the C-lobe of Ca^{2+}-CaM and RyR2. **g**, Local densities of the probable interacting residues, Tyr2157, Tyr2203, Arg2206 and Lys2154 in RyR2. The electron microscopy map was contoured at 5.5σ. **h**, Density of helix α−1 in the FKBP12.6/apo-CaM RyR2 structure. The sequence can be reliably assigned based on the indicated bulky residues. **i**, Density of helix α−1 in the FKBP12.6/ATP/caffeine/high-[Ca^{2+}]/Ca^{2+}-CaM RyR2 structure. The C-terminal half of helix α−1 is reliably assigned, a few bulky residues facilitate the sequence alignment. As both the N-terminal half of helix α−1 and C-lobe of Ca^{2+}-CaM had a lower resolution, the density shown here may belong to Trp3588. The electron microscopy maps in **h** and **i** were contoured at levels of 0.027 and 0.016, respectively.



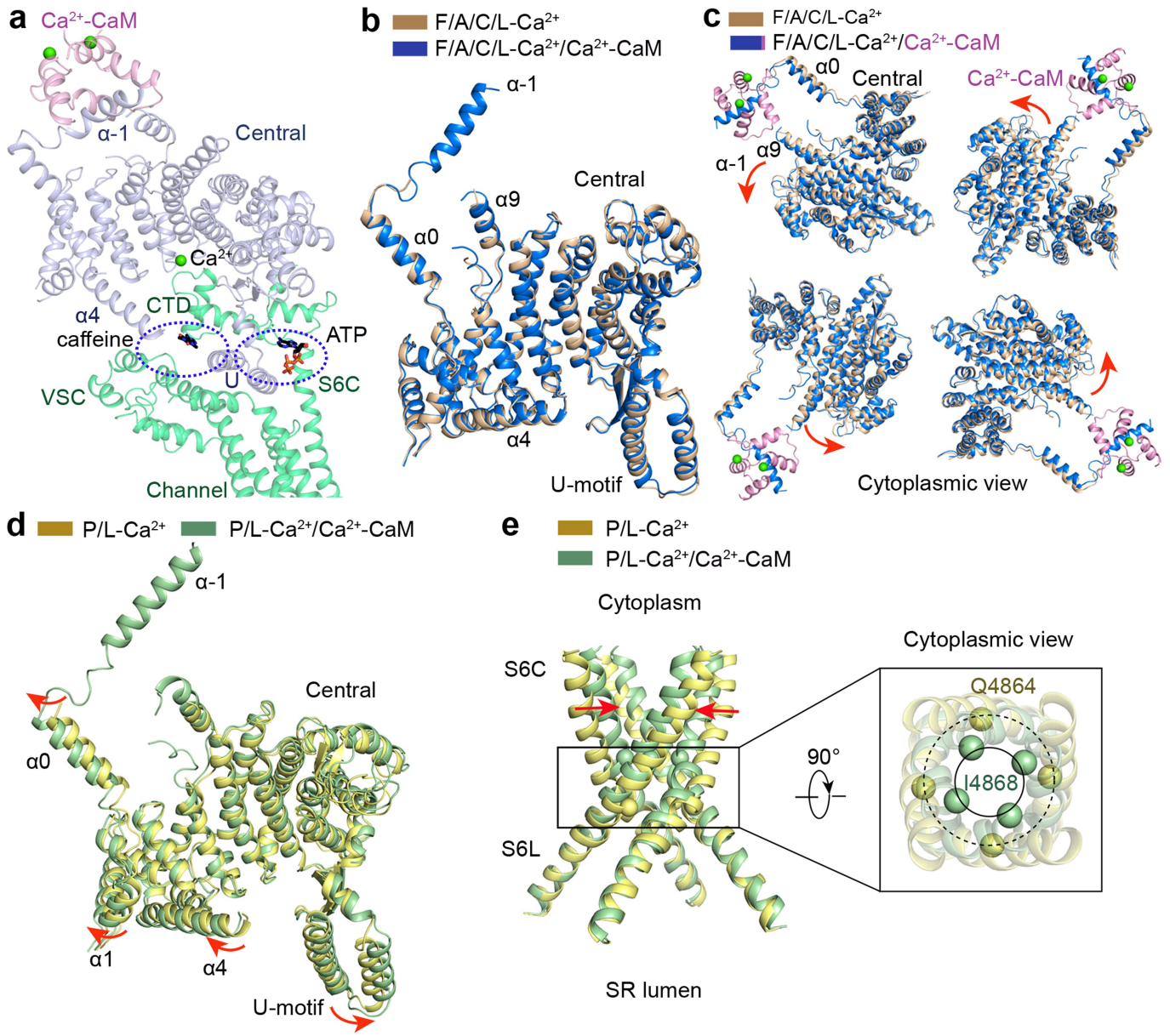
Extended Data Fig. 6 | Effect of RyR2 mutations on CaM regulation of single RyR2 channels. a–e, Single-channel activities were recorded in a symmetrical recording solution containing 250 mM KCl and 25 mM HEPES, pH 7.4. Representative current traces of single RyR2(WT) ($n = 9$), RyR2(Y2156A) ($n = 8$), RyR2(V3599A) ($n = 7$), RyR2(W3587A) ($n = 8$) and RyR2(L3590A) ($n = 9$) channels are shown. The Ca^{2+} concentration on the cytoplasmic and luminal face of the channel was 440 nM and around 45 nM, respectively. Open probability (Po), mean open time (To) and mean closed time (Tc) of the same channel before and after addition of CaM(WT) (1 μM) are depicted. Baselines are indicated by short bars on the right. **f**, Percentages of inhibition of channel open probability by CaM. Data are mean ± s.e.m. from single RyR2(WT) ($n = 9$), RyR2(Y2156A) ($n = 8$), RyR2(V3599A) ($n = 7$), RyR2(W3587A) ($n = 8$) and RyR2(L3590A) ($n = 9$) channels and analysed by one-way ANOVA with a Dunnett's post hoc test (versus RyR2(WT)) and adjusted *P* values are indicated on the graph.



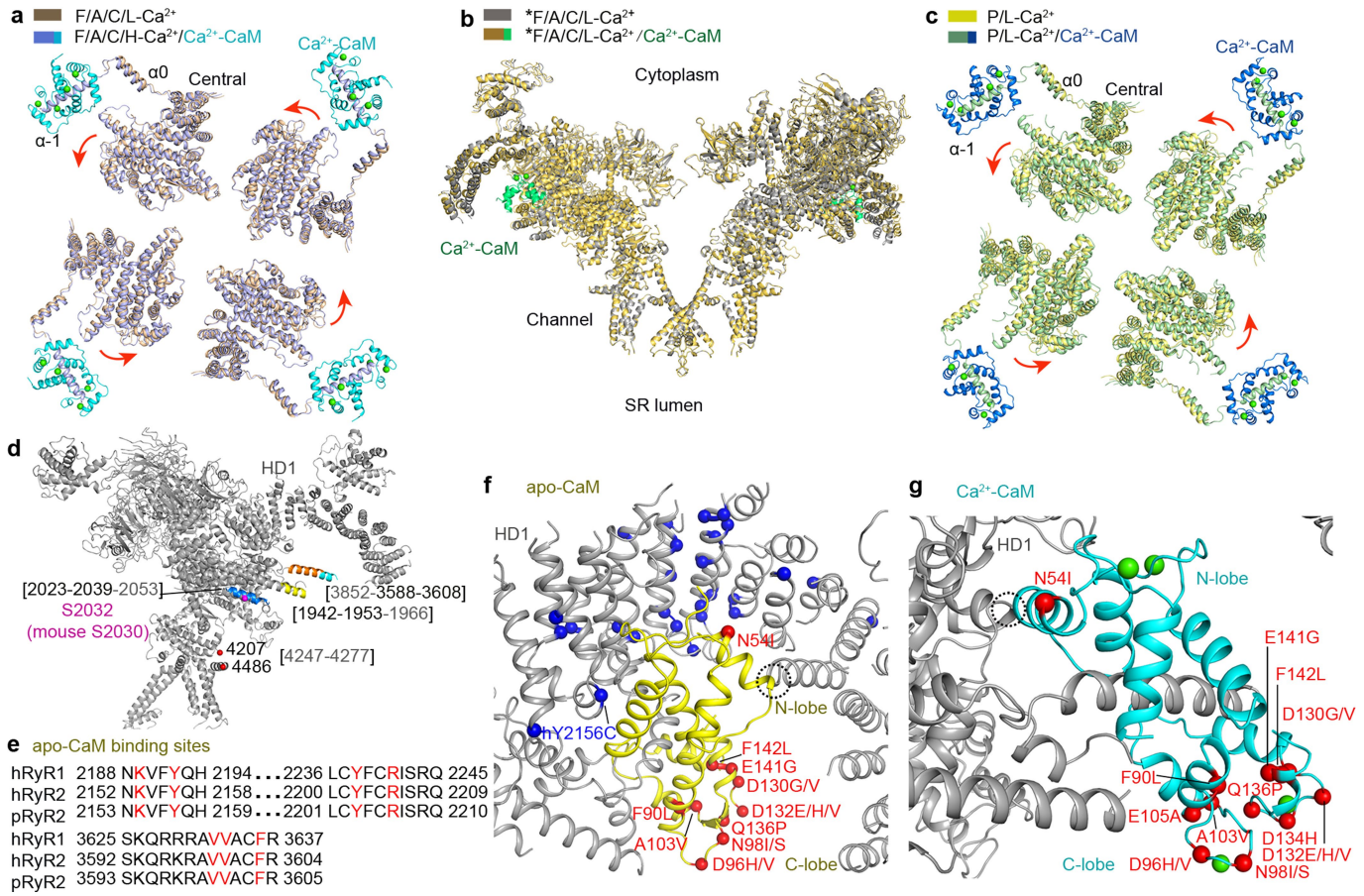
Extended Data Fig. 7 | Effect of RyR2 and CaM mutations on the termination of Ca^{2+} release in HEK293 cells. HEK293 cell lines that express RyR2(WT) and RyR2 mutants were co-transfected with the FRET-based endoplasmic-reticulum luminal Ca^{2+}-sensing protein D1ER and with no CaM (control), CaM(WT) or the Ca^{2+}-binding-deficient CaM mutant, CaM-M. **a**–**c**, Representative single-cell luminal Ca^{2+} recordings of RyR2(WT) cells transfected with no CaM (**a**; $n = 155$ cells), RyR2(WT) cells transfected with CaM(WT) (**b**; $n = 178$ cells) and RyR2-WT cells transfected with CaM-M (**c**; $n = 177$ cells) are shown. F_{SOICR} indicates the FRET level at which SOICR occurs and F_{termi} represents the FRET level at which SOICR terminates. The signal F_{max} is defined as the FRET level after tetracaine treatment. The minimum FRET signal maximum FRET F_{min} is defined as the FRET level after caffeine treatment. **d**, The activation threshold was determined as shown in **a**. **e**, The store capacity was calculated by subtracting F_{min} from F_{max}. **d**, **e**, Data are mean ± s.e.m. with the number of independent experiments for each condition shown on the graph and analysed by one-way ANOVA with a Dunnett's post hoc test (versus RyR2(WT) control) and adjusted P values are indicated. **f**–**h**, RyR2(WT) cells were co-transfected with the FRET-based endoplasmic-reticulum luminal Ca^{2+}-sensing protein D1ER and CaM(WT) or CaM mutants (CaM-M, CaM(E15A), CaM(F66A), CaM(L70A), CaM(M110A), CaM(F142A), CaM(F20A), CaM(F69A), CaM(F93A), CaM(L106A) and CaM(M146A)). CaM mutations close to a specific CaM–RyR2 interface are grouped and are indicated. The termination threshold (**f**), activation threshold (**g**) and store capacity (**h**) were determined as described in **a**–**e** above. **f**–**h**, Data are mean ± s.e.m. with the number of independent experiments for each condition shown on the graph and analysed by one-way ANOVA with a Dunnett's post hoc test (versus CaM(WT)) and adjusted P values are indicated.



Extended Data Fig. 8 | Evaluation of the conformations of N- and C-lobes of CaM in the FKBP12.6/ATP/caffeine/high-[Ca^{2+}]/Ca^{2+}-CaM structures. a, The electron microscopy map (low-pass-filtered to 4.8 Å resolution at a contour level of 0.015) of FKBP12.6/ATP/caffeine/high-[Ca^{2+}]/Ca^{2+}-CaM. Red and blue boxes indicate the N- and C-lobes, respectively. **b**, Docking of the three reported conformations of the N-lobe of CaM into our electron microscopy map suggests an open conformation of the N-lobe in FKBP12.6/ATP/caffeine/high-[Ca^{2+}]/Ca^{2+}-CaM. CC indicates the cross-correlation coefficient. **c**, An important distinction between C-lobes in the open and semi-open or closed states is the enlarged angle between helices C1 and C4. **d**, Docking analysis supports the open conformation of the C-lobe in FKBP12.6/ATP/caffeine/high-[Ca^{2+}]/Ca^{2+}-CaM.



Extended Data Fig. 9 | Inhibitory mechanism of RyR2 by Ca^{2+}-CaM. **a**, Both caffeine and ATP are located at the interfaces between the U-motif and O-ring, locking these elements into a stable unit. **b**, There is almost no intradomain rearrangement of the individual central domain between the structures of FKBP12.6/ATP/caffeine/low-[Ca^{2+}] and FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/Ca^{2+}-CaM. **c**, Ca^{2+}-CaM induces anticlockwise rotation of the overall central domain in FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/Ca^{2+}-CaM when viewed from the cytoplasmic side. The overall tetrameric FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/Ca^{2+}-CaM and FKBP12.6/ATP/caffeine/low-[Ca^{2+}] RyR2 structures are superimposed relative to C-terminal subdomain of the channel domain. Red arrows indicate the conformational changes upon Ca^{2+}-CaM binding. **d**, Ca^{2+}-CaM induces intradomain shifts of the individual central domain in the PCB95 and Ca^{2+}-activated channel. **e**, The PCB95 and Ca^{2+}-activated channel closes after Ca^{2+}-CaM loading.



Extended Data Fig. 10 | Conformational changes induced by Ca^{2+}-CaM and mapping of previously identified CaM-binding sequences and disease-associated point mutations onto the structures of RyR2–CaM complexes. a, Compared to FKBP12.6/ATP/caffeine/low-[Ca^{2+}], the four central domains in the FKBP12.6/ATP/caffeine/high-[Ca^{2+}]/Ca^{2+}-CaM RyR2 structure undergo an anticlockwise rotation. The overall tetrameric FKBP12.6/ATP/caffeine/low-[Ca^{2+}] and FKBP12.6/ATP/caffeine/high-[Ca^{2+}]/Ca^{2+}-CaM RyR2 structures are superimposed relative to the C-terminal subdomain of the channel domain. **b**, Compared to the CHAPS- and DOPC-treated FKBP12.6/ATP/caffeine/low-[Ca^{2+}] RyR2 structure, almost no conformational change was induced by Ca^{2+}-CaM in the CHAPS- and DOPC-treated FKBP12.6/ATP/caffeine/low-[Ca^{2+}]/Ca^{2+}-CaM RyR2 structure. **c**, Compared to the PCB95/low-[Ca^{2+}] RyR2 structure, the overall central domain in the PCB95/low-[Ca^{2+}]/Ca^{2+}-CaM RyR2 structure undergoes an anticlockwise rotation. The overall tetrameric PCB95/low-[Ca^{2+}] and PCB95/low-[Ca^{2+}]/Ca^{2+}-CaM RyR2 structures are superimposed relative to the C-terminal subdomain of the channel domain. **d**, Structural mapping of previously reported CaM-binding sequences. Orange, the overlapping binding sequences of apo-CaM and Ca^{2+}-CaM; cyan, the binding sequence of Ca^{2+}-CaM; yellow, the binding sequence of apo-CaM; blue, segments that are not involved in binding in our structures; red, sequences that are invisible in the structures. The residue numbers in brackets that are labelled grey indicate that the sequences are invisible in the structure. **e**, The primary apo-CaM binding sequences in RyR2 are the same in RyR1. Red residues highlight the key contact residues. **f**, Mapping of the disease-associated point mutations onto the structure of the RyR2–apo-CaM complex. The mutations in HD1 and apo-CaM are coloured blue and red, respectively. **g**, Mapping of the CaM disease-associated point mutations onto the structure of the RyR2–Ca^{2+}-CaM complex.

Extended Data Table 1 | Cryo-EM data collection, refinement and validation statistics

	F/apo-CaM (EMDB-9833) (PDB 6JI8)	F/A/C/L-Ca^{2+}/CaM-M (EMDB-9834) (PDB 6JII)	F/A/C/L-Ca^{2+} (EMDB-9831) (PDB 6JI0)	F/A/C/L-Ca^{2+}/Ca^{2+}-CaM (EMDB-9836) (PDB 6JIU)	*F/A/C/L-Ca^{2+} (EMDB-9879) (PDB 6JRR)	*F/A/C/L-Ca^{2+}/Ca^{2+}-CaM (EMDB-9880) (PDB 6JRS)	F/A/C/H-Ca^{2+}/Ca^{2+}-CaM (EMDB-9837) (PDB 6JIY)	P/L-Ca^{2+}/Ca^{2+}-CaM (EMDB-9889) (PDB 6JV2)
Data collection and processing								
Magnification	105,000	105,000	105,000	105,000	105,000	105,000	105,000	105,000
Voltage (kV)	300	300	300	300	300	300	300	300
Electron exposure (e–/Å^2)	45.6	50	50	45.6	50	50	50	50
Defocus range (μm)	-0.8~-1.8	-1.3~-1.7	-1.3~-1.7	-0.8~-1.8	-1.3~-1.7	-1.3~-1.7	-1.3~-1.7	-1.3~-1.7
Pixel size (Å)	1.338	1.091	1.091	1.338	1.091	1.091	1.091	1.091
Symmetry imposed	C4	C4	C4	C4	C4	C4	C4	C4
Initial particle images (no.)	1,180,104	404,978	515,941	1,026,058	769,264	666,134	595,005	759,504
Final particle images (no.)	208,715	69,556	78,841	77,029	149,212	154,111	96,158	22,876
Map resolution (Å)	3.6	4.2	4.2	4.2	3.9	3.7	3.9	4.4
FSC threshold	0.143	0.143	0.143	0.143	0.143	0.143	0.143	0.143
Map resolution range (Å)	535.2-3.6	436.4-4.2	436.4-4.2	535.2-4.2	436.4-3.9	436.4-3.7	436.4-3.9	436.4-4.4
Refinement								
Initial model used (PDB code)	5GOA	5GOA	5GOA	5GOA	5GOA	5GOA	5GOA	5GOA
Model resolution (Å)	3.6	4.2	4.2	4.2	3.9	3.7	3.9	4.4
FSC threshold	0.143	0.143	0.143	0.143	0.143	0.143	0.143	0.143
Model resolution range (Å)	535.2-3.6	436.4-4.2	436.4-4.2	535.2-4.2	436.4-3.9	436.4-3.7	436.4-3.9	436.4-4.4
Map sharpening *B* factor (Å^2)	-169	-167	-181	-167	-174	-161	-181	-64
Model composition								
Non-hydrogen atoms	115,060	115,028	109,772	112,212	109,132	112,208	115,288	111,080
Protein residues	15,024	15,012	14,332	14,684	14,268	14,660	15,056	14,564
Ligands	4	16	16	24	16	24	32	24
B factors (Å^2)								
Protein	79.52	192.63	147.15	226.13	118.38	103.61	122.58	303.05
Ligand	76.76	97.59	75.47	158.62	50.84	41.76	71.04	574.56
R.m.s. deviations								
Bond lengths (Å)	0.010	0.008	0.007	0.006	0.010	0.009	0.007	0.011
Bond angles (°)	1.250	1.096	1.123	0.981	1.111	1.050	1.033	1.292
Validation								
MolProbity score	2.06	2.01	2.03	2.11	2.02	2.08	2.04	2.63
Clashscore	8.56	7.61	7.43	10.69	8.45	8.70	7.78	19.3
Poor rotamers (%)	1.06	0.59	0.88	0.53	0.81	0.91	0.47	1.93
Ramachandran plot								
Favored (%)	89.0	88.4	86.9	89.5	89.5	87.3	87.5	86.5
Allowed (%)	10.5	11.2	12.9	10.3	10.2	12.4	12.2	13.3
Disallowed (%)	0.5	0.4	0.2	0.2	0.3	0.3	0.3	0.2

nature research

Corresponding author(s): Deshun Gong; S.R. Wayne Chen; Nieng Yan

Last updated by author(s): May 23, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	AutoEMationII; Compix Simple PCI 6; Axopatch 200A amplifier
Data analysis	Relion 2.0; Thunder; MotionCor2 1.1.0; GCTF 1.06; Phenix 1.12; Coot 0.8.6; Pymol 1.8.6.0; Chimera 1.11; pclamp 8.1; Compix Simple PCI 6, GraphPad Prism 8.1.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Atomic coordinates and EM density maps of the F/apo-CaM (PDB: 6JI8; EMDB: EMD-9833), F/A/C/L-Ca2+/CaM-M (PDB: 6JII; EMDB: EMD-9834), F/A/C/L-Ca2+ (PDB: 6JI0; EMDB: EMD-9831), F/A/C/L-Ca2+/Ca2+-CaM (PDB: 6JIU; EMDB: EMD-9836), *F/A/C/L-Ca2+ (PDB: 6JRR; EMDB: EMD-9879), *F/A/C/L-Ca2+/Ca2+-CaM (PDB: 6JRS; EMDB: EMD-9880), F/A/C/H-Ca2+/Ca2+-CaM (PDB: 6JIY; EMDB: EMD-9837), and P/L-Ca2+/Ca2+-CaM (PDB: 6JV2; EMDB: EMD-9889) structures have been deposited in the Protein Data Bank (http://www.rcsb.org) and the Electron Microscopy Data Bank (https://www.ebi.ac.uk/pdbe/emdb/). Source Data for Fig. 2e, f and Extended Data Figs. 1c, 6f, 7d-h are available online. All other data are available from the corresponding authors upon reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample size calculation was predetermined. We chose the sample size based on our previous studies that would be sufficient for testing for differences between groups.
Data exclusions	No data were excluded.
Replication	All experiments were repeated at least 5 times with similar results.
Randomization	No animals were involved in this study. Samples (cells or single channels) were not randomized for experiments in this study.
Blinding	Single channel recording and single cell calcium imaging experiments were not blinded, but the same recording and imaging conditions (solutions and recording/imaging settings) were applied to all channels and cell groups.

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	*Briefly describe the study type including whether data are quantitative, qualitative, or mixed-methods (e.g. qualitative cross-sectional, quantitative experimental, mixed-methods case study).*
Research sample	*State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving existing datasets, please describe the dataset and source.*
Sampling strategy	*Describe the sampling procedure (e.g. random, snowball, stratified, convenience). Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient. For qualitative data, please indicate whether data saturation was considered, and what criteria were used to decide that no further sampling was needed.*
Data collection	*Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper, computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection.*
Timing	*Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort.*
Data exclusions	*If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Non-participation	*State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation.*
Randomization	*If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled.*

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	*Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.*
Research sample	*Describe the research sample (e.g. a group of tagged Passer domesticus, all Stenocereus thurberi within Organ Pipe Cactus National Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets,*

describe the data and its source.

Sampling strategy	*Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.*
Data collection	*Describe the data collection procedure, including who recorded the data and how.*
Timing and spatial scale	*Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken*
Data exclusions	*If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Reproducibility	*Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.*
Randomization	*Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.*
Blinding	*Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.*

Did the study involve field work? ☐ Yes ☐ No

Field work, collection and transport

Field conditions	*Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).*
Location	*State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).*
Access and import/export	*Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).*
Disturbance	*Describe any disturbance caused by the study and how it was minimized.*

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	*Describe all antibodies used in the study; as applicable, provide supplier name, catalog number, clone name, and lot number.*
Validation	*Describe the validation of each primary antibody for the species and application, noting any validation statements on the manufacturer's website, relevant citations, antibody profiles in online databases, or data provided in the manuscript.*

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	The Flp-In T-REx HEK293 Cell Line was obtained from Invitrogen

Authentication	Not authenticated
Mycoplasma contamination	These cells tested negative for mycoplasma contamination
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used in this study.

Palaeontology

Specimen provenance	*Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information).*
Specimen deposition	*Indicate where the specimens have been deposited to permit free access by other researchers.*
Dating methods	*If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.*

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	The porcine hearts were bought from meat processing factory
Wild animals	NA
Field-collected samples	NA
Ethics oversight	NA

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics	*Describe the covariate-relevant population characteristics of the human research participants (e.g. age, gender, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."*
Recruitment	*Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.*
Ethics oversight	*Identify the organization(s) that approved the study protocol.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	*Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.*
Study protocol	*Note where the full trial protocol can be accessed OR if not available, explain why.*
Data collection	*Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.*
Outcomes	*Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.*

ChIP-seq

Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links
May remain private before publication.

For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.

Files in database submission

Provide a list of all files available in the database submission.

Genome browser session
(e.g. UCSC)

Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.

Methodology

Replicates

Describe the experimental replicates, specifying number, type and replicate agreement.

Sequencing depth

Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.

Antibodies

Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.

Peak calling parameters

Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.

Data quality

Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.

Software

Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.

Flow Cytometry

Plots

Confirm that:

☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.

Instrument

Identify the instrument used for data collection, specifying make and model number.

Software

Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.

Cell population abundance

Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.

Gating strategy

Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type

Indicate task or resting state; event-related or block design.

Design specifications

Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.

Behavioral performance measures

State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).

Acquisition

Imaging type(s) — *Specify: functional, structural, diffusion, perfusion.*

Field strength — *Specify in Tesla*

Sequence & imaging parameters — *Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.*

Area of acquisition — *State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.*

Diffusion MRI ☐ Used ☐ Not used

Preprocessing

Preprocessing software — *Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).*

Normalization — *If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.*

Normalization template — *Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.*

Noise and artifact removal — *Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).*

Volume censoring — *Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.*

Statistical modeling & inference

Model type and settings — *Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).*

Effect(s) tested — *Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.*

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference (See Eklund et al. 2016) — *Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.*

Correction — *Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).*

Models & analysis

n/a	Involved in the study
☐	☐ Functional and/or effective connectivity
☐	☐ Graph analysis
☐	☐ Multivariate modeling or predictive analysis

Functional and/or effective connectivity — *Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).*

Graph analysis — *Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).*

Multivariate modeling and predictive analysis — *Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.*

LETTER

https://doi.org/10.1038/s41586-019-1389-7

A fast radio burst localized to a massive galaxy

V. Ravi[1,2]*, M. Catha[3], L. D'Addario[1], S. G. Djorgovski[1], G. Hallinan[1], R. Hobbs[3], J. Kocz[1], S. R. Kulkarni[1], J. Shi[1], H. K. Vedantham[1,4], S. Weinreb[1] & D. P. Woody[3]

Intense, millisecond-duration bursts of radio waves (named fast radio bursts) have been detected from beyond the Milky Way[1]. Their dispersion measures—which are greater than would be expected if they had propagated only through the interstellar medium of the Milky Way—indicate extragalactic origins and imply contributions from the intergalactic medium and perhaps from other galaxies[2]. Although several theories exist regarding the sources of these fast radio bursts, their intensities, durations and temporal structures suggest coherent emission from highly magnetized plasma[3,4]. Two of these bursts have been observed to repeat[5,6], and one repeater (FRB 121102) has been localized to the largest star-forming region of a dwarf galaxy at a cosmological redshift of 0.19 (refs. [7–9]). However, the host galaxies and distances of the hitherto non-repeating fast radio bursts are yet to be identified. Unlike repeating sources, these events must be observed with an interferometer that has sufficient spatial resolution for arcsecond localization at the time of discovery. Here we report the localization of a fast radio burst (FRB 190523) to a few-arcsecond region containing a single massive galaxy at a redshift of 0.66. This galaxy is different from the host of FRB 121102, as it is a thousand times more massive, with a specific star-formation rate (the star-formation rate divided by the mass) a hundred times smaller.

We detected the fast radio burst (FRB) 190523 on 23 May 2019 (modified Julian date (MJD) 58626.254118233(2)), using the Deep Synoptic Array ten-antenna prototype (DSA-10; see Methods). (Throughout this paper, we quote standard errors (68% confidence limits) of the least-significant figures in parentheses.) The DSA-10 consists of 4.5-m radio dishes separated by 6.75 m to 1,300 m, located at the Owens Valley Radio Observatory. The DSA-10 is designed to detect FRBs in the phase-incoherent combination of signals from each dish, and then to process the same signals interferometrically (coherent combination) to localize FRBs to few-arcsecond accuracy. FRB 190523 was detected at a dispersion measure of 760.8(6) pc cm^{-3}, and localized to the following J2000 coordinates: right ascension (RA) 13 h 48 min 15.6(2) s; declination (dec.) +72° 28′ 11(2)″. A time-frequency dataset was formed at this position through the coherent addition of interferometric visibility data from DSA-10 (see Methods). These data, displayed in Fig. 1, consist of total-intensity spectra recorded in 1,250 frequency channels between 1,334.69 MHz and 1,487.28 MHz over 131.072-μs intervals, with the data in each channel incoherently corrected with at least 8.192-μs accuracy for the dispersive delay. The burst signal-to-noise ratio exceeds 10 in multiple time samples. The observed properties of FRB 190523 are summarized in Table 1. We derive a fluence of approximately 280 Jy ms given the sensitivity of DSA-10 at the burst location within the field of view. We detected no repeat bursts at this position during approximately 78 h of observations obtained over 54 days surrounding the detection (see Methods).

The 99% confidence containment region of FRB 190523 (Fig. 2) includes just one galaxy in archival data from the Panoramic Survey Telescope and Rapid Response System (Pan-STARRS) 3π Steradian Survey[10]. This galaxy, PSO J207.0643 + 72.4708 (hereafter PSO J207 + 72), was detected with an *r*-band magnitude of 22.1(1) in the stacked Pan-STARRS data. We obtained images of the containment region of the burst on MJD 58635 with the Keck I telescope of the W. M. Keck Observatory, using the Low Resolution Imaging Spectrometer (KeckI/LRIS; see Methods)[11]. We detected no objects other than PSO J207 + 72 within the FRB 190523 99% confidence containment region, to limiting magnitudes of 25.8 in the *g*-filter and 26.1 in the *R*-filter. The containment region lies within an apparent grouping of galaxies (Fig. 2), with the galaxy nearest to the containment region (S2 in Fig. 2) having been detected by Pan-STARRS with an *r*-band magnitude of 22.1(1). We also obtained a low-resolution optical spectrum of PSO J207 + 72 using KeckI/LRIS on MJD 58635 (see Methods).

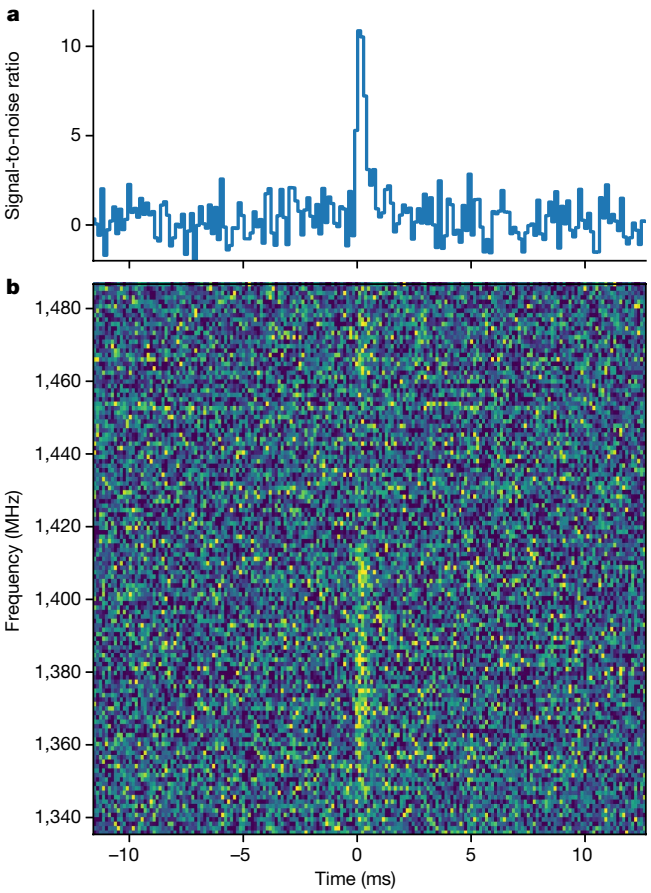


Fig. 1 | Time-frequency data on FRB 190523. a, Dedispersed temporal profile of the burst, averaged over the DSA-10 frequency band. The data are measures of the received power in 131.072-μs bins, in units of the root-mean-square (r.m.s.) off-burst signal-to-noise ratio. **b**, The dedispersed dynamic spectrum of the burst, again in units of the r.m.s. off-burst signal-to-noise ratio in each 1.22 MHz frequency channel. Although the structure evident in the burst spectrum is probably qualitatively accurate, no calibration of the relative flux density scales in different frequency channels has been applied.

[1]Cahill Center for Astronomy and Astrophysics, California Institute of Technology, Pasadena, CA, USA. [2]Center for Astrophysics, Harvard and Smithsonian, Cambridge, MA, USA. [3]Owens Valley Radio Observatory, California Institute of Technology, Big Pine, CA, USA. [4]ASTRON, Netherlands Institute for Radio Astronomy, Dwingeloo, The Netherlands. *e-mail: vikram@caltech.edu

Table 1 | Properties of FRB 190523 and its host galaxy

Property	Measurement
Topocentric arrival time at 1,530 MHz (MJD)	58626.254118233(2)
Fluence (Jy ms)	Greater than roughly 280
Dispersion measure (pc cm^{-3})	760.8(6)
Dispersion measure index	−2.003(8)
Milky Way disk (halo) dispersion measure (pc cm^{-3})	37 (50 to 80)
Extragalactic dispersion measure (pc cm^{-3})	644 to 674
Band averaged full-width half-maximum of the burst (ms)	0.42(5)
Scattering timescale at 1 GHz (ms)	1.4(2)
Right ascension (J2000)	13 h 48 min 15.6(2) s
Declination (J2000)	+72° 28′ 11(2)″
Host galaxy redshift	0.660(2)
Host galaxy luminosity distance (Gpc)	4.08(1)
Burst spectral energy (erg Hz^{-1})	5.6×10^{33}
Host galaxy stellar mass ($M_\odot$)	$10^{11.07(6)}$
Host galaxy star-formation rate ($M_\odot$ yr^{-1})	Less than 1.3(2)

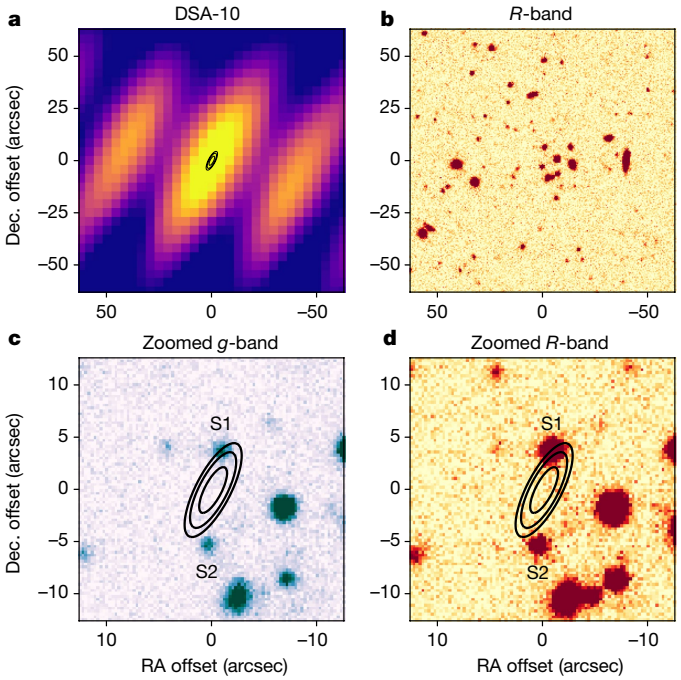


Fig. 2 | Images of the sky location of FRB 190523. All images are centred on J2000 coordinates RA 13 h 48 min 15.6(2) s; dec. +72° 28′ 11(2)″. **a**, Dirty snapshot image of the burst, obtained with DSA-10 (see Methods). **b**, Optical image in the *R*-band filter, obtained with KeckI/LRIS. The position of FRB 190523 coincides with an apparent grouping of galaxies. **c**, **d**, Zoom-in on the burst localization region in the *g*- and *R*-filters of KeckI/LRIS. The position of FRB 190523 is indicated with 68%, 95% and 99% confidence containment ellipses in **a**, **c**, **d**. The only galaxy detected above the 26.1-magnitude *R*-band detection limit within the 99% confidence containment ellipse, indicated by S1, is PSO J207+72. A galaxy to the south of the 99% confidence ellipse is labelled S2.

The spectrum (Fig. 3) indicates stellar absorption features at a redshift of 0.660(2). A single emission line, corresponding to the [O II] 3,727-Å doublet, is tentatively detected with a flux of $4.7(7) \times 10^{-17}$ erg s^{-1} cm^{-2}.

We modelled the Pan-STARRS optical photometry and the KeckI/LRIS optical spectroscopy of PSO J207 + 72 using the Prospector software[12,13]. We used this software to fit a 'delay-tau' stellar population and star-formation-history model to the data. In this model, the star-formation history is proportional to $te^{-t/\tau}$, where t is the time since the formation epoch of the galaxy, and τ is the characteristic decay time of the star-formation history. We derive a metallicity fraction of 0.3(2) of the Solar metallicity, a stellar mass of $10^{11.07(6)} M_\odot$ (where $M_\odot$ is the mass of the Sun), and an ongoing star-formation rate of approximately $1.3 M_\odot$ yr^{-1}. The star-formation rate—although poorly constrained given the limited wavelength coverage of the data—is consistent with the flux of the possible [O II] 3,727-Å emission line, which implies a star-formation rate of up to $1.3(2) M_\odot$ yr^{-1} (ref. [14]). As this emission line could also be associated with weak activity of the central black hole[15], we adopt a star-formation rate of $1.3(2) M_\odot$ yr^{-1} as an approximate upper limit for PSO J207 + 72.

As the only object detected within the containment region of FRB 190523, PSO J207 + 72 is the likely host galaxy of the burst. Additional evidence is furnished by the agreement between the burst dispersion measure and the predicted dispersion measure for the redshift of PSO J207 + 72. Accounting for 37 pc cm^{-3} from the Milky Way disk[16], and between 50 pc cm^{-3} and 80 pc cm^{-3} from the Milky Way ionized halo[17], the extragalactic dispersion measure of FRB 190523 is between 644 pc cm^{-3} and 674 pc cm^{-3}. Given this extragalactic dispersion measure, and parameterizing the containment region by the 3 × 8 arcsec 95% confidence ellipse[18], the probability of finding any galaxy (even one not detectable in our data) by chance within the containment region is less than 10% (ref. [18]). Further, the redshift of PSO J207 + 72 is not larger than would be expected given the dispersion measure of FRB 190523. The dispersion measure contributed by the intergalactic medium (IGM) to the redshift of PSO J207 + 72 is $660(f_{\rm IGM}/0.7)$ pc cm^{-3}, where $f_{\rm IGM}$ is the fraction of the luminous matter (termed baryons) of the Universe in the ionized IGM[19]. Observations suggest that 60% of cosmic baryons are in the IGM ($f_{\rm IGM}$ equals approximately 0.6(1)), 10% of baryons are locked in galaxies, and the remaining 30% of baryons are apportioned between the circumgalactic medium in galaxy halos and the IGM[20]. We adopt $f_{\rm IGM} = 0.7$ as a fiducial value, noting that a root-mean-squared scatter of roughly 200 pc cm^{-3} in the IGM dispersion measure to redshifts of around 0.66 is expected owing to cosmic variance and intervening galaxy halos[21]. Finally, our *R*-band KeckI/LRIS image excludes the possibility that the FRB 190523 containment region includes a dwarf galaxy like the host of the repeating FRB 121102 at a redshift below approximately 0.45 (corresponding to a luminosity distance of 2.5 gigaparsecs)[9]. The less than 10% probability of chance coincidence of the burst containment region with a galaxy—even one as small as the FRB 121102 host—implies that there is a less than 10% probability that the FRB 190523 containment region includes a galaxy like the FRB 121102 host. This further suggests that PSO J207 + 72 is the unique host galaxy of FRB 190523.

The properties of FRB 190523 are typical of FRBs observed at frequencies of around 1.4 GHz (ref. [22]). At the distance of PSO J207 + 72, FRB 190523 has a spectral energy of 5.6×10^{33} erg Hz^{-1}, which is consistent with the largest previously estimated burst energies[23]. The patchy spectrum of FRB 190523 (Fig. 1) is also similar to the spectra of bright FRBs detected by the Australian Square Kilometre Array Pathfinder[23]. We note that our DSA-10 observations cannot exclude the possibility of repeated bursts from FRB 190523 below our detection threshold, which FRB 190523 itself exceeded by only 15%.

The temporal profile of FRB 190523 indicates a broadening timescale, owing to multipath propagation through inhomogeneous plasma, of 1.4(2) ms at 1 GHz (ref. [22]). This broadening timescale is three orders of magnitude higher than expected for the sightline of FRB 190523 through the Milky Way interstellar medium[16]. The broadening timescale is also larger than would be expected during propagation through the dispersion-measure column potentially contributed by PSO J207 + 72 (less than roughly 150 pc cm^{-3})[24]. Our results therefore support the possibility of some FRBs being temporally broadened during propagation between their host galaxies and the Milky Way, for example, in the circumgalactic medium of intervening galaxies[25].

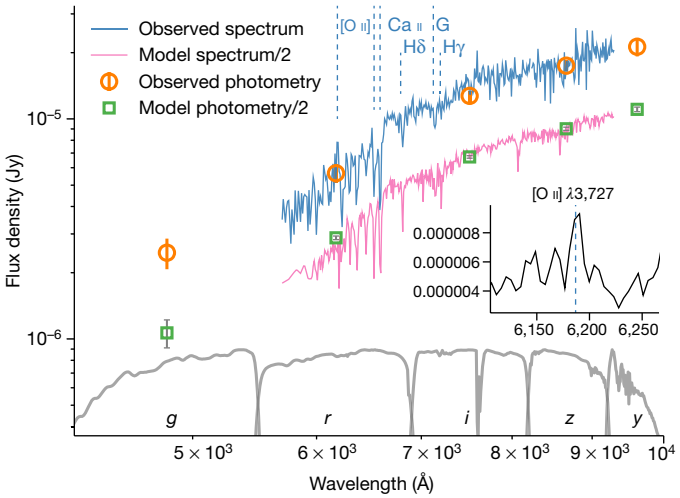


Fig. 3 | Modelling of the host galaxy of FRB 190523. We obtained a low-resolution optical spectrum (blue line) of PSO J207 + 72 using KeckI/LRIS on MJD 58635 (see Methods). We also modelled the Pan-STARRS optical photometry (orange circles) and the KeckI/LRIS optical spectroscopy of PSO J207 + 72 using Prospector software[12,13] (pink line). Error bars denoting one standard deviation are shown for the Pan-STARRS photometry. The maximum a posteriori probability (MAP) results from the Prospector modelling of the host galaxy are scaled downwards by a factor of two. The grey curves illustrate transmissions from the Pan-STARRS *g*-, *r*-, *i*-, *z*- and *y*-filters. The grey error bars accompanying the MAP photometry points (green boxes) indicate the 5th and 95th percentiles of 500 samples drawn from the posterior parameter distributions. The redshifted positions of some notable absorption lines are indicated by dashed blue traces. The inset shows the observed spectrum around the [O II] 3,727-Å feature, binned by a factor of two less than the spectrum in the main panel.

The properties of PSO J207 + 72 are in tension with FRB progenitor models developed on the basis of the host galaxy of the repeating FRB 121102 (ref. [26]). In particular, the host of FRB 121102 is similar to the dwarf-star-forming host galaxies of superluminous supernovae and long gamma-ray bursts, which are the terminal explosions of the most massive stars. However, the stellar mass of PSO J207 + 72 is higher and its star-formation rate per unit mass is lower than those of the known host galaxies of superluminous supernovae and long gamma-ray bursts at redshifts below 1 (ref. [26]). In addition, leading models for the FRB emission mechanism favour neutron-star progenitors with magnetar magnetic-field strengths (of greater than roughly 10^{14} G)[3,4,27]. If this is the case, then our results suggest that magnetars that were formed in the terminal explosions of the most massive stars are not the only objects capable of emitting FRBs. Indeed, magnetar-formation channels exist that do not require young stellar populations, such as the accretion-induced collapse of white dwarfs to neutron stars in mass-transfer binaries[28,29], and the merger of two neutron stars[30].

The likely low contribution of PSO J207 + 72 to the dispersion measure of FRB 190523 provides evidence in support of FRB progenitor models (magnetar or otherwise) that do not require actively star-forming environments. The low global star-formation rate of PSO J207 + 72, together with the spatially offset location of much of the containment region of FRB 190523 relative to the galaxy (Fig. 2), leads us to consider the possibility that the progenitor of FRB 190523 was drawn from an old stellar population. The similarity between the stellar populations of PSO J207 + 72 and the Milky Way suggests that galaxies like the Milky Way can harbour FRB progenitors.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1389-7.

Received: 11 June 2019; Accepted: 25 June 2019;
Published online 2 July 2019.

1. Lorimer, D. R., Bailes, M., McLaughlin, M. A., Narkevic, D. J. & Crawford, F. A bright millisecond radio burst of extragalactic origin. *Science* **318**, 777–780 (2007).
2. Petroff, E., Hessels, J. W. T. & Lorimer, D. R. Fast radio bursts. *Astron. Astrophys. Rev.* **27**, 4 (2019).
3. Kumar, P., Lu, W. & Bhattacharya, M. Fast radio burst source properties and curvature radiation model. *Mon. Not. R. Astron. Soc.* **468**, 2726–2739 (2017).
4. Metzger, B. D., Margalit, B. & Sironi, L. Fast radio bursts as synchrotron maser emission from decelerating relativistic blast waves. *Mon. Not. R. Astron. Soc.* **485**, 4091–4106 (2019).
5. Spitler, L. et al. A repeating fast radio burst. *Nature* **531**, 202–205 (2016).
6. The CHIME/FRB Collaboration. A second source of repeating fast radio bursts. *Nature* **566**, 235–238 (2019).
7. Chatterjee, S. et al. A direct localization of a fast radio burst and its host. *Nature* **541**, 58–61 (2017).
8. Marcote, B. et al. The repeating fast radio burst FRB 121102 as seen on milliarcsecond angular scales. *Astrophys. J.* **834**, L8 (2017).
9. Bassa, C. et al. FRB 121102 is coincident with a star-forming region in its host galaxy. *Astrophys. J.* **843**, L8 (2017).
10. Chambers, K. C. et al. The Pan-STARRS1 surveys. Preprint at https://arxiv.org/abs/1612.05560 (2016).
11. Oke, J. B. et al. The Keck low-resolution imaging spectrometer. *Publ. Astron. Soc. Pacif.* **107**, 375–385 (1995).
12. Leja, J. & Johnson, B. D. bd-j/prospector: initial release. *Zenodo* https://doi.org/10.5281/zenodo.1116491 (2017).
13. Leja, J., Johnson, B. D., Conroy, C., van Dokkum, P. G. & Byler, N. Deriving physical properties from broadband photometry with rrospector: description of the model and a demonstration of its accuracy using 129 galaxies in the local Universe. *Astrophys. J.* **837**, 170 (2017).
14. Rosa-González, D., Terlevich, E. & Terlevich, R. An empirical calibration of star formation rate estimators. *Mon. Not. R. Astron. Soc.* **332**, 283–295 (2002).
15. Yan, R. et al. On the origin of [O II] emission in red-sequence and poststarburst galaxies. *Astrophys. J.* **648**, 281–298 (2006).
16. Cordes, J. M. & Lazio, T. J. W. NE2001. I. A new model for the galactic distribution of free electrons and its fluctuations. Preprint at http://arxiv.org/abs/astroph/0207156 (2002).
17. Prochaska, J. X. & Zheng, Y. Probing galactic haloes with fast radio bursts. *Mon. Not. R. Astron. Soc.* **485**, 648–665 (2019).
18. Eftekhari, T. & Berger, E. Associating fast radio bursts with their host galaxies. *Astrophys. J.* **849**, 162 (2017).
19. Shull, J. M. & Danforth, C. W. The dispersion of fast radio bursts from a structured intergalactic medium at redshifts $z < 1.5$. *Astrophys. J.* **852**, L11 (2018).
20. Shull, J. M., Smith, B. D. & Danforth, C. W. The baryon census in a multiphase intergalactic medium: 30% of the baryons may still be missing. *Astrophys. J.* **759**, 23 (2012).
21. McQuinn, M. Locating the "missing" baryons with extragalactic dispersion measure estimates. *Astrophys. J.* **780**, L33 (2014).
22. Ravi, V. The observed properties of fast radio bursts. *Mon. Not. R. Astron. Soc.* **482**, 1966–1978 (2019).
23. Shannon, R. M. et al. The dispersion-brightness relation for fast radio bursts from a wide-field survey. *Nature* **562**, 386–390 (2018).
24. Bhat, N. D. R., Cordes, J. M., Camilo, F., Nice, D. J. & Lorimer, D. R. Multifrequency observations of radio pulse broadening and constraints on interstellar electron density microstructure. *Astrophys. J.* **605**, 759–783 (2004).
25. Vedantham, H. K. & Phinney, E. S. Radio wave scattering by circumgalactic cool gas clumps. *Mon. Not. R. Astron. Soc.* **483**, 971–984 (2019).
26. Metzger, B. D., Berger, E. & Margalit, B. Millisecond magnetar birth connects FRB 121102 to superluminous supernovae and long-duration gamma-ray bursts. *Astrophys. J.* **841**, 14 (2017).
27. Lyutikov, M. Coherence constraints on physical parameters at bright radio sources and FRB emission mechanism. Preprint at https://arxiv.org/abs/1901.03260 (2019).
28. Piro, A. L. & Kollmeier, J. A. Ultrahigh-energy cosmic rays from the "en caul" birth of magnetars. *Astrophys. J.* **826**, 97 (2016).
29. Ruiter, A. J. et al. On the formation of neutron stars via accretion-induced collapse in binaries. *Mon. Not. R. Astron. Soc.* **484**, 698–711 (2019).
30. Giacomazzo, B. & Perna, R. Formation of stable magnetars from binary neutron star mergers. *Astrophys. J.* **771**, L26 (2013).

METHODS

The DSA-10 instrument. The Deep Synoptic Array 10-element prototype (DSA-10) is an array of ten 4.5-m radio dishes operating in the frequency band 1.28–1.53 GHz. The array is deployed at the Owens Valley Radio Observatory (OVRO; located at 37.2314 °N, 118.2941 °W) near the town of Bishop, California, USA. A description of the DSA-10 instrument is given in ref. [31]. Here we describe the state of the instrument at the time that FRB 190523 was detected.

The array was in a slightly modified configuration relative to its initial deployment, with four antennas clustered at the northern end of the OVRO T-shaped infrastructure. The positions of each antenna, in standard International Terrestrial Reference Frame (ITRF) geocentric coordinates, are given in Extended Data Table 1. Each antenna was equipped with two receivers sensitive to orthogonal linear polarizations. The antenna primary beams have full-width half-maxima of 3.25°. Antenna 2 was not operational because it was being used to test new equipment, and one polarization of antenna 8 was operating with substantially reduced sensitivity caused by a malfunctioning low-noise amplifier. Antenna 2 was discarded from all calibration and imaging procedures described below. The array operated in a stationary drift-scan mode on the meridian at a declination of +73.6°, with an absolute pointing accuracy of better than 0.4°. The projected baseline lengths ranged between 5.75 m and 1,256.57 m.

The DSA-10 was operated in this configuration between MJD 58568 and MJD 58630, with a total time on-sky of 54 days. FRB searching was conducted using the incoherent sum of dynamic spectra from the eight fully functioning antennas, forming a single stream of 2,048-channel spectra integrated over 131.072 μs. Before summation, the dynamic spectra were excised of narrowband and impulsive broadband radio-frequency interference (RFI)[31]. We searched these data for FRBs in real time using the Heimdall software[32], with 2,477 optimally spaced dispersion-measure trials between 30 pc cm^{-3} and 3,000 pc cm^{-3}. At each trial dispersion measure, the data were smoothed with boxcar filters spaced by powers of two between 2^0 and 2^8 samples before searching. The detection threshold was set at eight standard deviations (8σ). In this study, we assume a typical band-averaged system-equivalent flux density of 22 kJy for each DSA-10 receiver, on the basis of interferometric measurements of the system sensitivity using sources with known flux densities[31]. Given eight fully functioning antennas, and 220 MHz of effective bandwidth following RFI excision, this implies an approximate detection threshold of 94 Jy ms at the centre of the primary beam for a millisecond-duration FRB not affected by intrachannel dispersion smearing[33].

Upon detection of any pulse candidate that exceeded the detection threshold at any trial dispersion measure, 294,912 samples of complex voltage data corresponding to each polarization of each antenna were written to disk. These data consisted of 4-bit real, 4-bit imaginary 2,048-channel voltage spectra sampled every 8.192 μs, calculated on and transmitted to five servers by Smart Network ADC Processor (SNAP-1) boards[34] over a 10-gigabit ethernet network. The data dumps were extracted from ring buffers such that the candidate pulse arrival times at 1,530 MHz were 61,035 samples into the dumps.

These voltage data were also used to derive interferometric visibilities between each pair of antennas. The visibilities were measured by integrating the cross-power over 0.402653184 s, and over 625 pairs of channels between exactly 1,334.6875 MHz and 1,487.275390625 MHz. Approximate, constant-path-length delay corrections were digitally applied to each receiver input on the SNAP-1 boards, but no time-dependent fringe-tracking corrections were applied online. Visibility data were recorded only when bright unresolved radio sources were transiting through the DSA-10 primary beam. These data were fringe-stopped offline by dividing the data by a model for the visibilities given the known source positions from the National Radio Astronomy Observatory (NRAO) Very Large Array (VLA) Sky Survey (NVSS) catalogue[35]. Visibility modelling was accomplished using differential antenna positions referenced to the known ITRF location of the centre of the OVRO T-shape (which had previously hosted the Caltech OVRO Millimeter Array), using the Common Astronomy Software Applications (CASA, version 5.1.1) package to calculate baseline coordinates. We consider visibility data on three such sources here: NVSS J120019 + 730045 (also 3C 268.1, 5.56 Jy; hereafter J1200 + 7300), NVSS J145907 + 714019 (7.47 Jy; hereafter J1459 + 7140) and NVSS J192748 +735802 (3.95 Jy; hereafter J1927 + 7358). Data on these sources were recorded for 3,630 s, 1,960 s and 3,890 s, respectively, centred on their transit times.

Interferometric calibration and localization of FRB 190523. We used standard strategies for processing radio-interferometric data[36] to calibrate the instrumental responses of each DSA-10 antenna and receiver. Here we describe the specific methods used to calibrate the data on FRB 190523, and the steps taken to verify their efficacy.

FRB 190523 was detected on MJD 58626.254118233(2), and a voltage-data dump was successfully triggered. These data were cross-correlated offline using the same routines as applied in the online correlator software[37], and the visibilities were integrated over 131.072 μs. Only data in 1,250 channels covering the frequency band (1,334.6875–1,487.275390625 MHz) spanned by the visibility data recorded in real time were retained in the analysis presented here.

At the time FRB 190523 was detected, the DSA-10 pointing centre was at a position (J2000) of RA 14 h 15 min 01.98 s, dec. +73° 40′ (absolute pointing accuracy of better than 0.4°). The calibrator sources J1459 + 7140 and J1200 + 7300 transited 29.58 min later and 163.76 min earlier, respectively. The phases of the per-receiver complex gain corrections for the FRB 190523 data were derived as follows. No attempt at per-receiver gain amplitude calibration was made. This was because all sources under consideration (including FRB 190523) were consistent with unresolved point sources, based on NVSS data[35], that dominated the sky brightness within their fields. All visibility amplitudes were taken to be unity, such that only phase information was preserved.

First, receiver-based relative delay errors (with antenna 7 as a reference) were calculated using fringe-stopped data on J1459 + 7140, restricted to the 15 min surrounding transit. J1459 + 7140 is considered to be a primary calibrator in the database of the VLA for baseline lengths consistent with the DSA-10.

Second, after applying these delay corrections to the 15 min of J1459 + 7140 data surrounding transit, the data were averaged in time, and in frequency to 25 channels. The averaged data were used to derive receiver-based phase errors in each channel.

Third, the phase solutions from J1459 + 7140 were averaged with phase solutions derived from 15 min of fringe-stopped data on J1200 + 7300 surrounding transit, with the same delay corrections as above applied first. No substantial differences were evident between the phase solutions derived independently from J1459 + 7140 and J1200 + 7300.

Fourth, the delay and phase solutions from the above analysis were used to calibrate the visibility data on FRB 190523. The phase centre was set to the array pointing centre at the time of the burst. The data were converted to the measurement-set format for further analysis with CASA. Data on the four shortest baselines (after removing baselines with antenna 2) were excluded because of substantial levels of correlated noise. A 7° × 7° total-intensity image, without deconvolution of the synthesized beam shape (a 'dirty' image), was then made using four visibility time-samples centred on the burst, with the standard imaging task tclean applied for gridding and Fourier inversion. A single point-like source was evident in this image, at a position 2.3° from the pointing centre (an hour angle of 26.8′ west and 1.2° south).

Fifth, given the apparent offset location of the burst from the pointing centre, we then corrected for any direction-dependent instrumental-response variations intrinsic to the DSA-10 antennas. This was done by extracting 6 min of fringe-stopped data on J1200 + 7300 at the same hour angle as the possible position of FRB 190523, applying the previous calibration solutions, and deriving frequency-averaged phase corrections for each receiver (again using antenna 7 as a reference). We note that no data on J1459 + 7140 were available at the hour angle of the burst, as visibilities were recorded on this source for a shorter time (1,960 s) than for J1200 + 7300 (3,630 s). Large corrections of up to 25° in phase were required, which were identical for the two receivers on each antenna. This formed the final set of calibration solutions for FRB 190523.

We then applied these final calibration solutions to the visibility data on FRB 190523, and referenced the data to a phase centre corresponding to the approximate burst position, with the burst dispersion accounted for in calculating baseline coordinates. The data were then summed over the two polarizations, and converted to the measurement-set format. The CASA task tclean was used to make dirty and deconvolved images of the burst data (see Fig. 2 and the bottom row of Extended Data Fig. 1). The imaging process was verified using the 6 min of data on J1200 + 7300 obtained at the same hour angle as FRB 190523 (Extended Data Fig. 1, top row). No sources were detected in images made using visibility data in 128 time samples on either side of FRB 190523, either when averaged together or when binned by four samples.

The position of FRB 190523 was estimated by fitting to the calibrated visibilities, using four 131.072-μs time samples centred on the burst, as before. This fit was carried out using the MIRIAD[38] task uvfit (after converting the measurement set to a MIRIAD-format file) and a grid-search code, as a software problem in the CASA task uvmodelfit prevented it from loading our data. The grid-search code was used to evaluate the posterior probability of the source position given the likelihood of the visibility data over a uniform 0.25-arcsec grid of positions centred on the burst position in its image. This was then used to calculate the maximum a posteriori probability location quoted in Table 1, and the 68%, 95% and 99% confidence containment ellipses shown in Fig. 2. We also attempted to estimate the position of FRB 190523 using only data between 1,350 MHz and 1,420 MHz, where much of the burst spectral energy density appears to be concentrated (Fig. 1). This yielded a containment ellipse that was consistent with the result from all the data, but with major and minor axes that were 10% larger.

We verified the efficacy of our localization procedure using a selection of methods. First, no substantial phase closure errors were evident in the calibrated data on J1200 + 7300 and J1459 + 7140, either at boresight, or at the hour angle of FRB 190523 in the case of J1200 + 7300. No baseline-based calibration corrections were required to accurately model the calibrator data. Second, we verified that the calibration solutions derived as above for FRB 190523 were also able to calibrate visibility data on the source J1927 + 7358, which transited five hours after the burst detection. We did this by extracting 6 min of data on J1927 + 7358 at the hour angle that the burst was detected, applying the same calibration solutions as applied to the burst, and imaging it and deriving its position as for the burst (Extended Data Fig. 1, middle row). The position of J1927 + 7358 was recovered to within 1 arcsec in both dimensions, with the offsets consistent with the position-fit errors. Plots of the calibrated, frequency-averaged visibility phases on each baseline of 6 min of data on J1200 + 7300 and J1927 + 7358 after rotation to their known positions are shown in Extended Data Fig. 2, together with the same results for FRB 190523 rotated to its derived position. The single worst outliers in Extended Data Fig. 2 for FRB 190523 and J1927 + 7358 were both on baselines containing antenna 1.

We repeated the same calibration procedure as above on the 12 days of data before detection of the burst, and correctly recovered the position of J1927 + 7358 on each day (Extended Data Fig. 3). The r.m.s. scatter in the recovered positions of J1927 + 7358 about the true value was 0.47 arcsec in RA, and 0.69 arcsec in dec. We therefore have no basis to add a systematic-error contribution to the position-fit errors for FRB 190523. We have previously verified that no temporal error existed in the voltage-data dumps by imaging giant pulses from the Crab pulsar (B0531 + 21) when the DSA-10 was pointed at its declination, running the same software[31]. We ensured that this remained the case by calibrating and imaging data dumps obtained close to when J1200 + 7300 was transiting on other days using the above procedures, and verifying that the position of J1200 + 7300 was correctly recovered.

The final astrometric reference for our results was the VLA calibrator catalogue, which was accurate to less than 0.01 arcsec for J1459 + 7140 and J1927 + 7358, and to the NVSS accuracy of approximately 0.5 arcsec (ref. [35]) for J1200 + 7300. These errors are small in comparison with the final localization accuracy of FRB 190523, and hence we do not include them in the localization-error budget of the burst.

Properties of FRB 190523. We modelled the temporal profile of FRB 190523 using published methods[22]. The data presented in Fig. 1 were formed by the coherent addition of calibrated visibility data on FRB 190523 using its best-fit position. These data were integrated over five evenly spaced frequency bands, and the resulting time series were fit with a series of models. The best-fit model was the convolution of the instrumental response to a delta-function impulse and a one-sided exponential with a timescale varying as f^{-4}, where f is the observed frequency. This is consistent with temporal broadening caused by multipath propagation. The extrapolated broadening timescale at 1 GHz is quoted in Table 1. We also quote the uncertainty in the dispersion-measure index in Table 1; we found the burst arrival time to scale with $f^{-2.003(8)}$.

We made no attempt to calibrate the response of the DSA-10 to polarized radiation. The DSA-10 was not designed for polarimetry. First, we have not established our ability to robustly calibrate the per-receiver frequency-dependent gain amplitudes using transiting continuum calibrator sources. We also do not record full-polarization visibility data on these sources, making it impossible to measure signal leakages between the receivers that are sensitive to orthogonal linear polarizations. The lack of polarization information is not likely to affect the burst localization, because each polarization was calibrated independently using unpolarized sources. We verified that consistent positions for FRB 190523 were derived from data in each polarization separately.

KeckI/LRIS observations and analysis. KeckI/LRIS observations of the localization region and candidate host galaxy of FRB 190523 (PSO J207 + 72) were carried out on the night of MJD 58635 in dark time, under clear photometric conditions with a median seeing-disk full-width half-maximum (R-band) of 1.1 arcsec. Light from the telescope was split between the two arms of LRIS by the D560 dichroic. Three images were taken in the g- and R-band filters at an airmass of 1.66, with exposure times of 30 s, 300 s and 300 s, and no binning of the detector pixels. The g- and R-band filters have effective wavelengths of 4,731 Å and 6,417 Å, respectively, and effective widths of 1,082 Å and 1,185 Å, respectively. Three spectral exposures were obtained (with exposure times of 900 s and a median airmass of 1.68) with a 1.5-arcsec long slit at a position angle of 270°, the 600/4,000 grism for the blue arm, the 400/8,500 grating for the red arm, and the detector binned by two pixels in the spectral direction. The spectral-flux calibration was obtained with observations of the standard star Feige 67 at an airmass of 1.07.

All optical data were reduced using standard procedures for LRIS. Bias subtraction using the overscan levels, flat-fielding using dome-flat exposures, and cosmic-ray rejection were performed with lpipe software[39]. The imaging data were then astrometrically registered against Gaia data-release 2 (DR2) stars using scamp software routines[40], co-added using swarp software routines[41], and sources were extracted using the SExtractor software[42]. Photometric calibration to an accuracy of 0.1 magnitudes was accomplished using Pan-STARRS objects in the field. The weakest detected sources in the g- and R-bands were 25.8 and 26.1 magnitudes (AB) respectively, which we adopt as our limiting magnitudes in these bands.

We also used the lpipe software to process the spectroscopic data by performing wavelength calibration using internal-arc exposures corrected by sky-emission lines, and by optimally subtracting the sky-emission lines. We then performed optimal extraction of the spectral traces in each on-source exposure by using the trace of the standard star Feige 67, which we also used for flux calibration and the removal of telluric absorption lines. The final optimally co-added spectrum of PSO J207 + 72 has a flux calibration uncertainty of 10% owing to the differing air masses of the standard-star and source observations. The galaxy was detected only in the red arm of LRIS, and a truncated spectrum (displayed in Fig. 3) was used for further analysis. In addition to the [O II] 3,727-Å emission-line doublet, some other detected absorption lines (Ca II H and K lines at 3,935 Å and 3,970 Å respectively, Hγ at 4,342 Å and Hδ at 4,103 Å, and the Fraunhofer G feature at 4,306 Å) are labelled in Fig. 3. All lines were detected at a redshift of 0.660(2).

Modelling of the host galaxy. We modelled the Pan-STARRS photometry and KeckI/LRIS spectrum of PSO J207 + 72 using the Prospector code for stellar-population inference. Prospector enables Markov Chain Monte Carlo (MCMC) sampling of the posterior distribution of parameters of the stellar populations and star-formation histories of galaxies, given a combination of photometric and spectroscopic data. Galaxy emission is modelled using a wrapper to the flexible stellar-population synthesis code[43,44]. We fit a five-parameter 'delay-tau' model for the stellar population of PSO J207 + 72, including the metallicity, the stellar-population age and star-formation timescale, the mass in formed stars, and the V-band extinction of a dust screen. Before performing the fit, we corrected the observations for Galactic extinction using the 'extinction' software package[45], through a standard Milky Way extinction curve with a V-band extinction of 0.052 magnitudes. Data surrounding the detected [O II] 3,727-Å emission-line doublet were masked, and no modelling of nebular emission was conducted. We conducted exploration of the posterior parameter distributions using the emcee MCMC software[46]. Standard Prospector priors were implemented. We derived a metallicity of 0.3(2) of the solar metallicity, a mass in formed stars of $10^{11.07(6)} M_{\odot}$, an age of 6.6(8) Gyr, a star-formation timescale of 1.0(2) Gyr, and a V-band extinction of 0.3(2) magnitudes.

PSO J207 + 72 lies within what appears to be a group of galaxies (Fig. 2) with Pan-STARRS r-band magnitudes ranging between 19 and 23. No spectra are available at present for galaxies within this group, and the association in distance cannot therefore be confirmed. PSO J207 + 72 is undetected in observations from the VLA Sky Survey[47]; the upper limit at 3 GHz on any source within the 99% confidence containment region of FRB 190523 is 0.36 mJy (3σ). Throughout this paper, we use cosmological parameters from the 2015 Planck analysis[48].

Data availability

The datasets generated during and/or analysed during this study are available from the corresponding author on reasonable request.

Code availability

Custom code is made available at https://github.com/VR-DSA.

31. Kocz, J. et al. DSA-10: a prototype array for localizing fast radio bursts. Preprint at https://arxiv.org/abs/1906.08699 (2019).
32. Barsdell, B. R., Bailes, M., Barnes, D. G. & Fluke, C. J. Accelerating incoherent dedispersion. *Mon. Not. R. Astron. Soc.* **422**, 379–392 (2012).
33. Cordes, J. M. & McLaughlin, M. A. Searches for fast radio transients. *Astrophys. J.* **596**, 1142–1154 (2003).
34. Hickish, J. et al. A decade of developing radio-astronomy instrumentation using CASPER open-source technology. *J. Astron. Instrum.* **5**, 1641001 (2016).
35. Condon, J. J. et al. The NRAO VLA Sky Survey. *Astron. J.* **115**, 1693–1716 (1998).
36. Thompson, A. R., Moran, J. M. & Swenson, G. W. Jr *Interferometry and Synthesis in Radio Astronomy* 3rd edn (Springer, 2017).
37. Clark, M. A., LaPlante, P. C. & Greenhill, L. J. Accelerating radio astronomy cross-correlation with graphics processing units. *Int. J. High Perform. Comput. Appl.* **27**, 178–192 (2013).
38. Sault, R. J., Teuben, P. J. & Wright, M. C. H. A retrospective view of MIRIAD. In *Astronomical Data Analysis Software and Systems IV, ASP Conf. Ser.* vol. 77 (eds Shaw, R. A et al.) 433 (1995).
39. Perley, D. A. Fully-automated reduction of longslit spectroscopy with the low resolution imaging spectrometer at Keck Observatory. Preprint at https://arxiv.org/abs/1903.07629 (2019).
40. Bertin, E. Automatic astrometric and photometric calibration with SCAMP. In *Astronomical Data Analysis Software and Systems XV, ASP Conf. Ser.* vol. 351 (eds Gabriel, C. et al.) 112 (2006).
41. Bertin, E. et al. The TERAPIX pipeline. In *Astronomical Data Analysis Software and Systems XI, ASP Conf. Proc.* vol. 281 (eds Bohlender, D. A. et al.) 228 (2002).
42. Bertin, E. & Arnouts, S. SExtractor: software for source extraction. *Astron. Astrophys. Suppl. Ser.* **117**, 393–404 (1996).

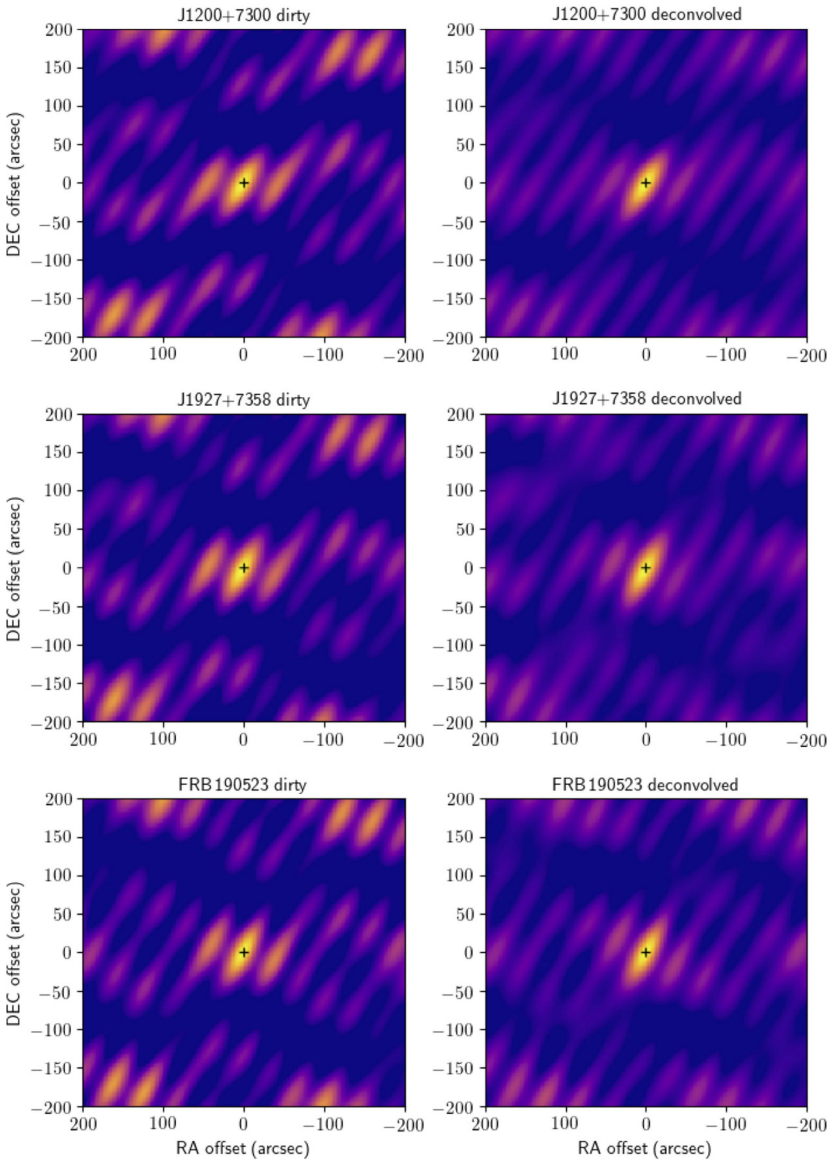
43. Conroy, C., Gunn, J. E. & White, M. The propagation of uncertainties in stellar population synthesis modeling. I. The relevance of uncertain aspects of stellar evolution and the initial mass function to the derived physical properties of galaxies. *Astrophys. J.* **699**, 486–506 (2009).
44. Conroy, C. & Gunn, J. E. The propagation of uncertainties in stellar population synthesis modeling. III. Model calibration, comparison, and evaluation. *Astrophys. J.* **712**, 833–857 (2010).
45. Barbary, K. extinction v0.3.0. *Zenodo* https://doi.org/10.5281/zenodo.804967 (2016).
46. Foreman-Mackey, D., Hogg, D. W., Lang, D. & Goodman, J. emcee: the MCMC hammer. *Publ. Astron. Soc. Pacif.* **125**, 306 (2013).
47. Murphy, D. & Lacy, M. VLA sky survey. https://science.nrao.edu/science/surveys/vlass (2019).
48. Planck Collaboration. Planck 2015 results. XIII. Cosmological parameters. *Astron. Astrophys.* **594**, A13 (2016).

Acknowledgements We thank the staff of the Owens Valley Radio Observatory, including J. Lamb, K. Hudson, A. Rizo and M. Virgin, for their assistance with the construction of the DSA-10. We thank A. Readhead for supporting the initiation of the DSA-10 project. We also thank A. Soliman for assistance with the development of the DSA-10 receivers. A portion of this research was performed at the Jet Propulsion Laboratory, California Institute of Technology, under a President and Directors Fund grant and under a contract with the National Aeronautics and Space Administration (NASA). This research was additionally supported by the National Science Foundation (NSF) under grant AST-1836018. V.R. acknowledges support as a Millikan Postdoctoral Scholar in Astronomy at the California Institute of Technology, and from a Clay Postdoctoral Fellowship of the Smithsonian Astrophysical Observatory. S.G.D. acknowledges partial support from NSF grant AST-1815034 and NASA grant 16-ADAP16-0232. Some of the data presented herein were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California and NASA. The Observatory was made possible by the financial support of the W. M. Keck Foundation. This research made use of Astropy, a community-developed core Python package for astronomy. The Pan-STARRS1 Surveys (PS1) and the PS1 public science archive have been made possible through contributions by the Institute for Astronomy, the University of Hawaii, the Pan-STARRS Project Office, the Max-Planck Society and its participating institutes, the Max-Planck Institute for Astronomy, Heidelberg, the Max-Planck Institute for Extraterrestrial Physics, Garching, Johns Hopkins University, Durham University, the University of Edinburgh, Queen's University Belfast, the Harvard-Smithsonian Center for Astrophysics, the Las Cumbres Observatory Global Telescope Network Incorporated, the National Central University of Taiwan, the Space Telescope Science Institute, NASA under grant NNX08AR22G issued through the Planetary Science Division of the NASA Science Mission Directorate, NSF grant AST-1238877, the University of Maryland, Eotvos Lorand University, the Los Alamos National Laboratory, and the Gordon and Betty Moore Foundation.

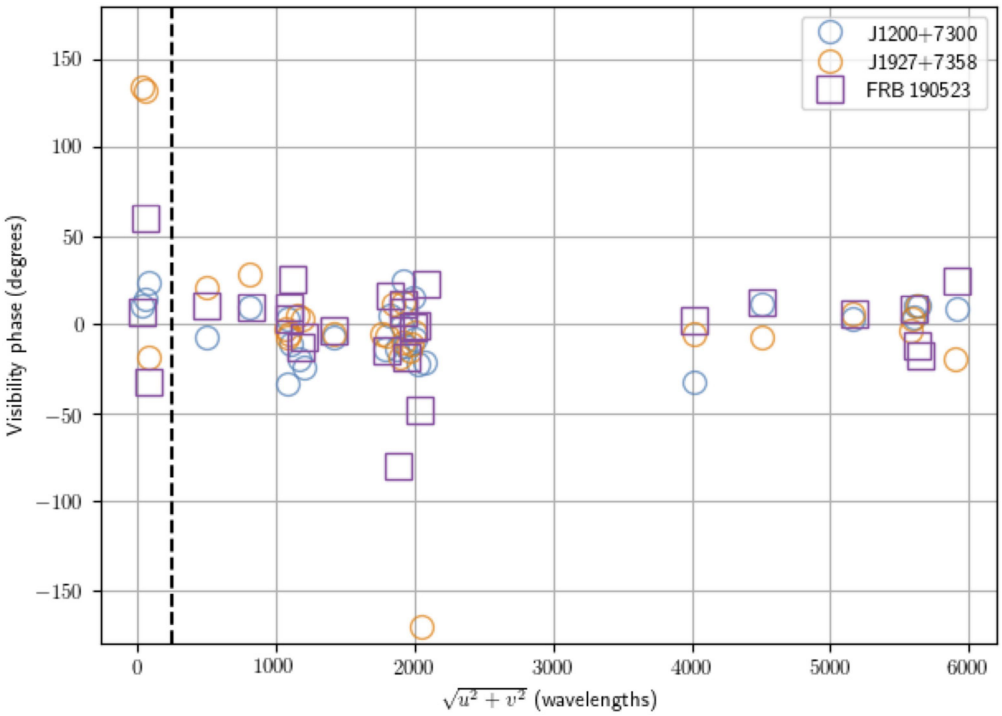
Author contributions G.H., V.R. and H.K.V. conceived of and developed the DSA-10 concept and observing strategy. V.R., J.K. and S.R.K. led the construction and initial deployment of DSA-10. D.P.W., S.W, L.D., J.K., V.R., H.K.V, M.C., R.H. and J.S. designed and built the DSA-10 subsystems. V.R. and H.K.V. commissioned the DSA-10. V.R. operated the DSA-10 and analysed the data. S.G.D. carried out the optical observations. V.R. analysed the optical data, and led the writing of the manuscript with the assistance of all co-authors.

Competing interests : The authors declare no competing interests.

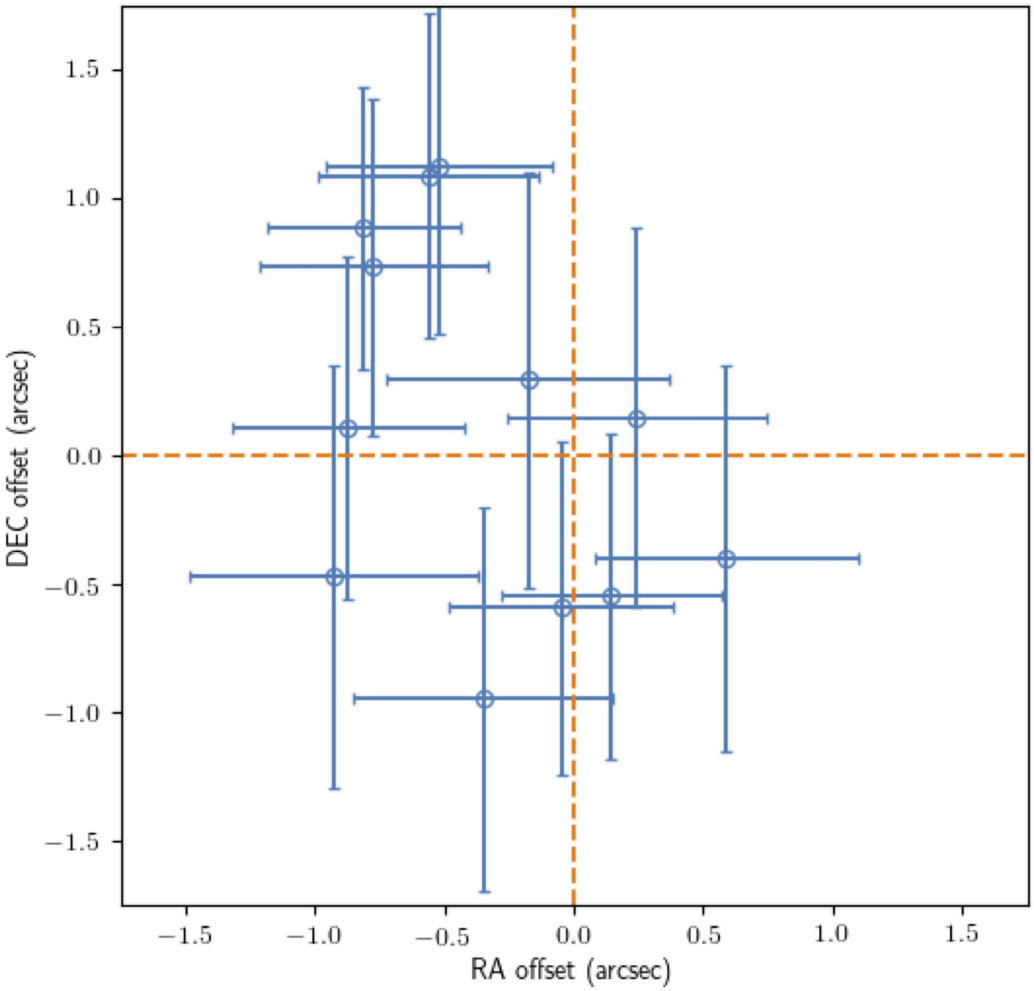
Additional information
Correspondence and requests for materials should be addressed to V.R.
Peer review information *Nature* thanks Shami Chatterjee and Jason Hessels for their contribution to the peer review of this work.



Extended Data Fig. 1 | DSA-10 images. Dirty and deconvolved images are shown of two bright point-sources and FRB 190523. All data were obtained at the same hour angle relative to the meridian, within 12 h of each other. The same calibration solution, derived using the J1200 + 7300 data, was applied to all data. The black crosses indicate the known source positions in the top and middle rows, and the best-fit position of FRB 190523 in the bottom row. The recovery of the correct position of J1927 + 7358 at the hour angle that FRB 190523 was detected at demonstrates the accuracy of the calibration solutions.



Extended Data Fig. 2 | Visibility phases measured for two bright point-sources and FRB 190523. Only data on baselines including fully functioning antennas are shown. The visibility data were phase-rotated to the known (or best-fit for FRB 190523) source positions, and averaged across the frequency band. Data on the shortest baselines (to the left of the dashed vertical line) were corrupted by correlated noise, and were discarded from imaging analysis. All data were calibrated using the same calibration solution, which was partially based on the J1200 + 7300 data, and were obtained at the same hour angle relative to the meridian within a 12-h timeframe. The *x* axis shows the baseline lengths in units of wavelengths at the middle of the DSA-10 frequency band.



Extended Data Fig. 3 | Recovered positions of J1927 + 7358 on 12 separate days. Each position was derived from 5 min of visibility data, extracted when J1927 + 7358 was at the same hour angle as FRB 190523 was detected. On each day, the data were also calibrated in exactly the same way as the FRB 190523 data. The error bars indicate the 68% (1σ) confidence intervals for the estimated positions.

Extended Data Table 1 | ITRF coordinates for the ten DSA-10 antennas

Antenna	X (m)	Y (m)	Z (m)
1	-2409464.509	-4477971.270	3839125.031
2	-2409466.445	-4477974.866	3839119.657
3	-2409470.315	-4477982.059	3839108.908
4	-2409547.552	-4478125.603	3838894.418
5	-2409468.38	-4477978.463	3839114.282
6	-2409429.957	-4478294.47	3838772.061
7	-2409682.474	-4478158.598	3838772.061
8	-2409746.758	-4478124.008	3838772.061
9	-2409770.667	-4478111.143	3838772.061
10	-2410525.007	-4477850.573	3838597.062

LETTER

https://doi.org/10.1038/s41586-019-1470-2

The formation of Jupiter's diluted core by a giant impact

Shang-Fei Liu[1,2]*, Yasunori Hori[3,4], Simon Müller[5], Xiaochen Zheng[6,7], Ravit Helled[5], Doug Lin[8,9] & Andrea Isella[2]

The Juno mission[1] has provided an accurate determination of Jupiter's gravitational field[2], which has been used to obtain information about the planet's composition and internal structure. Several models of Jupiter's structure that fit the probe's data suggest that the planet has a diluted core, with a total heavy-element mass ranging from ten to a few tens of Earth masses (about 5 to 15 per cent of the Jovian mass), and that heavy elements (elements other than hydrogen and helium) are distributed within a region extending to nearly half of Jupiter's radius[3,4]. Planet-formation models indicate that most heavy elements are accreted during the early stages of a planet's formation to create a relatively compact core[5–7] and that almost no solids are accreted during subsequent runaway gas accretion[8–10]. Jupiter's diluted core, combined with its possible high heavy-element enrichment, thus challenges standard planet-formation theory. A possible explanation is erosion of the initially compact heavy-element core, but the efficiency of such erosion is uncertain and depends on both the immiscibility of heavy materials in metallic hydrogen and on convective mixing as the planet evolves[11,12]. Another mechanism that can explain this structure is planetesimal enrichment and vaporization[13–15] during the formation process, although relevant models typically cannot produce an extended diluted core. Here we show that a sufficiently energetic head-on collision (giant impact) between a large planetary embryo and the proto-Jupiter could have shattered its primordial compact core and mixed the heavy elements with the inner envelope. Models of such a scenario lead to an internal structure that is consistent with a diluted core, persisting over billions of years. We suggest that collisions were common in the young Solar system and that a similar event may have also occurred for Saturn, contributing to the structural differences between Jupiter and Saturn[16–18].

Giant impacts[19,20] are likely to occur shortly after runaway gas accretion when Jupiter's gravitational perturbation increases to about thirty-fold in a fraction of a million years, thus destabilizing the orbits of nearby planetary embryos. This transition follows oligarchic growth[21] and the emergence of multiple embryos with isolation mass in excess of a few Earth masses, $M_\oplus$ (ref. [22]). Some of these massive embryos may collide with the gas giant during their orbit crossing[23,24]. Through tens of thousands of gravitational N-body simulations with different initial conditions, such as Jupiter's growth model, orbital configuration, and so on (see Methods), we find that the emerging Jupiter had a strong influence on nearby planetary embryos. As a result, in a large fraction of these numerical tests an embryo could collide with Jupiter within a few million years, that is, within the lifetime of the Solar nebula. Of those catastrophic events, head-on collisions are more common than grazing ones owing to Jupiter's gravitational focusing effects.

To investigate the influence of such impacts on the internal structure of the young Jupiter we use the hydrodynamics code FLASH[25] with the relevant equation of state (EOS). Details of the computational setup and the simulations are presented in Methods. In general, the disintegration of the intruding embryo leads to the disruption of the planet's original core. However, to establish a large diluted-core structure—as has been inferred from recent Jupiter structure models based on the Juno mission's measurements—the heavy-element material of the core and of the embryo need to mix efficiently with the surrounding gaseous envelope, which requires a large embryo to strike the young Jupiter almost head-on. Massive embryos are available at this early stage of the Solar System and our N-body simulations also suggest that head-on collisions are common (see Methods).

In Fig. 1 we show the consequence of a head-on collision between an embryo and Jupiter with an initial silicate+ice core mass of $M_{\rm core} = 10M_\oplus$, a hydrogen+helium (H-He) envelope, an approximately present-day total mass and radius (the young Jupiter may have been up to twice its present-day size, but, to avoid introducing additional free parameters, we consider models in which Jupiter is closer to its present-day size). In fact, the post-impact core–envelope structure depends mainly on the mass of the initial core and envelope as well as the impactor's mass and impact velocity $V_{\rm imp}$. We adopt an impact speed of $V_{\rm imp} \approx 46$ km s^{-1}, which is close to the free-fall speed onto Jupiter's surface (see Methods) and we assume that the impactor is comprised of an $8M_\oplus$ silicate+ice core and a $2M_\oplus$ H-He envelope. The combined total mass of the core of the proto-Jupiter and the core of the embryo, $M_{Z,\rm total}$, is chosen to be compatible with the mass of heavy elements (Z) derived from internal structure models of Jupiter with a diluted core[3]. We note that at Jupiter's distance of 5.2 astronomical units (AU) from the Sun, the impactor's speed relative to the gas giants is limited by the planets' surface escape speed. The acquisition of planetary embryos would not lead to any major changes in the spin angular momentum and orientation of the targeted planet. The total energy injected into the young Jupiter by the intruding embryo is only a few per cent of its original value so that there is little change in Jupiter's mean density and mass.

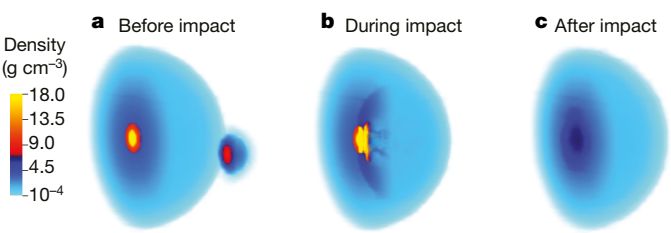


Fig. 1 | Three-dimensional cutaway snapshots of density distributions during a merger event between a proto-Jupiter with a $10M_\oplus$ rock/ice core and a $10M_\oplus$ impactor. a, Just before the contact. **b**, The moment of core–impactor contact. **c**, 10 h after the merger. Owing to impact-induced turbulent mixing, the density of Jupiter's core decreases by a factor of three after the merger, resulting in an extended diluted core. A two-dimensional presentation of density slices of the same event is shown in Extended Data Fig. 3.

[1]School of Physics and Astronomy, Sun Yat-sen University, Zhuhai, China. [2]Department of Physics and Astronomy, Rice University, Houston, TX, USA. [3]Astrobiology Center, Tokyo, Japan. [4]National Astronomical Observatory of Japan, Tokyo, Japan. [5]Institute for Computational Science, Center for Theoretical Astrophysics and Cosmology, University of Zurich, Zurich, Switzerland. [6]Department of Astronomy, Tsinghua University, Beijing, China. [7]Department of Physics, Tsinghua University, Beijing, China. [8]Department of Astronomy and Astrophysics, University of California, Santa Cruz, Santa Cruz, CA, USA. [9]Institute for Advanced Study, Tsinghua University, Beijing, China. *e-mail: liushangfei@mail.sysu.edu.cn

Table 1 | Initial conditions and final outcomes of the head-on giant impact simulation

	M_J	$M_{Z,core}$	$M_{impactor}$	$M_{Z,impactor}$	$M_{Z,total}$	R_{core}/R_J
Before merger	306.714	9.962	9.967	7.975	17.937	0.15
About 10 h after merger	304.946/313.360	17.693	–	–	17.901/17.925	0.423
H-4.5 (after 4.56 Gyr)	313.36	10.61	–	–	17.925	0.30
H-radenv (after 4.56 Gyr)	313.36	17.24	–	–	17.925	0.39
H-4.5-rock (after 4.56 Gyr)	313.36	15.92	–	–	17.925	0.45

M_J is the mass of the proto-Jupiter and $M_{impactor}$ is the mass of the impactor. $M_{Z,core}$ is the mass of heavy elements (silicate+ice) in the proto-Jupiter's core, and $M_{Z,impactor}$ is the total mass of heavy elements contained within the impactor. $M_{Z,total}$ is the total mass of heavy elements contained within the impactor and Jupiter. R_{core}/R_J is the radius of the proto-Jupiter's core scaled to Jupiter's present-day radius. Before the impact, the proto-Jupiter's core is completely made of heavy elements. $M_{Z,core}$ equals the core mass. After the impact, the proto-Jupiter's core is diluted with H and He. The new boundary of the diluted core is defined at the location where the mass fraction of heavy elements Z drops below 0.014, and $M_{Z,core}$ then equals the mass of a diluted core excluding H and He. Because the proto-Jupiter expands substantially after the merger (see Extended Data Fig. 3d), the values of the total mass of Jupiter (M_J) and the mass of heavy elements within Jupiter ($M_{Z,total}$) are measured within $1R_J$ and $2R_J$, respectively. Those values reveal that the majority of Jupiter's mass still resides within its original size, although a hot, extended, low-density envelope mostly made of H-He forms immediately after the merger. The last three rows list values for the evolution models that best fit an interior structure of Jupiter containing a diluted core[3]. All mass quantities are in units of $M_\oplus$.

The impact results in little mass loss (see Table 1), but Jupiter's initial core is completely disrupted. During the impactor's plunge towards and collision with the primordial core, a large amount of kinetic energy is dissipated. Heat release near the centre of Jupiter increases the local temperature T, offsets the pressure P balance, and induces oscillations (see the full animation in Supplementary Information). The inner part of the envelope becomes convective, driven by the steep temperature gradient near the core. Vigorous turbulence stirs up efficient mixing between the heavy elements and the H-He envelope. After a few dynamical timescales (a characteristic time with which to measure the expansion or contraction of a planet; Jupiter's dynamical timescale is a bit less than half an hour), the initial silicate+ice core is thoroughly mixed with the surrounding H-He envelope and their mass fraction is $Z \leq 0.5$ within 20% of Jupiter's radius R_J. Within about 30 dynamical timescales, Jupiter's interior settles into a quasi-hydrodynamic equilibrium with a diluted core extending to a radius of $0.4R_J$–$0.5R_J$ (see Table 1, Fig. 2a). In the outer half of the envelope, the gas density is slightly elevated and a small trace of the dredged-up heavy elements (silicate+ice) leads to the formation of a composition gradient.

The post-impact heavy-element distribution leads to a composition gradient that could evolve and become similar to an internal structure that has a diluted core. However, the hydrodynamic simulation is terminated ten hours after the impact. To explore under what conditions a diluted-core-like structure persists after the 4.56 Gyr of Jupiter's evolution, we compute the thermal evolution from shortly after the impact until the present day. The hydro-simulation sets the initial heavy-element gradient as shown in Fig. 2a. Because the exact post-impact temperature profile is unknown (it depends on the formation process[26,27], the energetics of the impact, and other such factors), we consider various temperature profiles with different central temperatures. Furthermore, we consider an initial thermal structure that accounts for the accretion shock during runaway gas accretion as suggested by a recent Jupiter formation model[27] (see Methods for details). We find that for the head-on collision, a post-impact central temperature of around 30,000 K leads to a present-day Jupiter with a diluted core. If the initial temperature profile is shaped by the accretion shock, this provides another model pathway to a diluted core for Jupiter. In Fig. 2b we show the density profiles of the best-fitting models after the 4.56 Gyr of evolution. If the central temperatures are higher (for example, 50,000 K), the interior is hot enough to 'delete' the heavy-element gradient, leading to a fully mixed planetary interior. On the other hand, for low central temperatures (about 20,000 K), convective mixing is less efficient and the inferred density profile is less consistent with a diluted-core structure. Therefore, we conclude that Jupiter's diluted-core structure could be explained by a giant impact event, but only under specific conditions including a head-on collision with a massive planetary embryo, a post-impact central temperature of about 30,000 K or an initial thermal structure created by the accretion shock during the runaway phase. Indeed, the hydrodynamic simulation suggests that most of the impact energy is not deposited in the deep interior, and therefore the central temperature is unlikely to increase substantially, supporting the diluted core solution (see Methods).

In contrast, if the same embryo collides with Jupiter at a grazing angle, it would be gradually tidally disrupted while sinking towards the centre of Jupiter (see Fig. 3). In Methods, we further show that impactors with one Earth mass (or less) disintegrate in the envelope of a gas giant before reaching its centre. Without smashing into the core directly, the shock wave induced by the impactor alone is insufficient to dredge up heavy elements from the core into Jupiter's envelope. Such impacts generally lead to core growth rather than core destruction. Since impacts of planetary embryos are expected to be frequent after a gas giant's runaway gas accretion phase, such an event with different impact conditions (such as a small impactor or an oblique collision) may have also happened to Saturn, and could in principle explain the differences between the internal structures of Jupiter and Saturn[16–19].

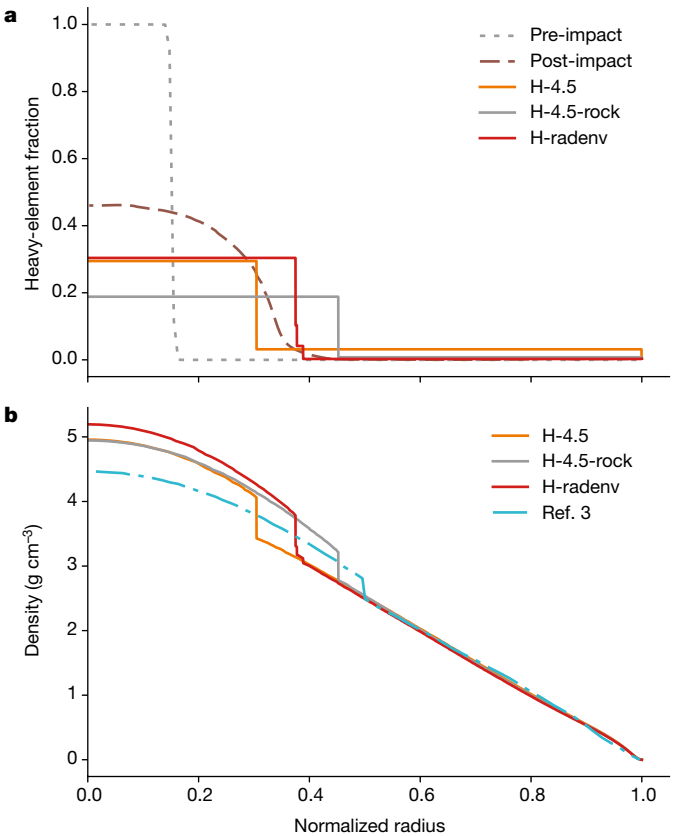


Fig. 2 | Post-impact thermal evolution models. a, Heavy-element distribution versus normalized radius before (dotted line) and after (dashed line) the giant impact. The solid lines show the composition after 4.56 Gyr of evolution for the three best-fit models that result in a diluted core; see Table 1, Methods and Extended Data Table 2 for more details. **b**, Density versus normalized radius after 4.56 Gyr of evolution for three best-fit models (solid lines) and from the diluted-core interior structure model of Wahl et al.[3] (dash-dotted line).

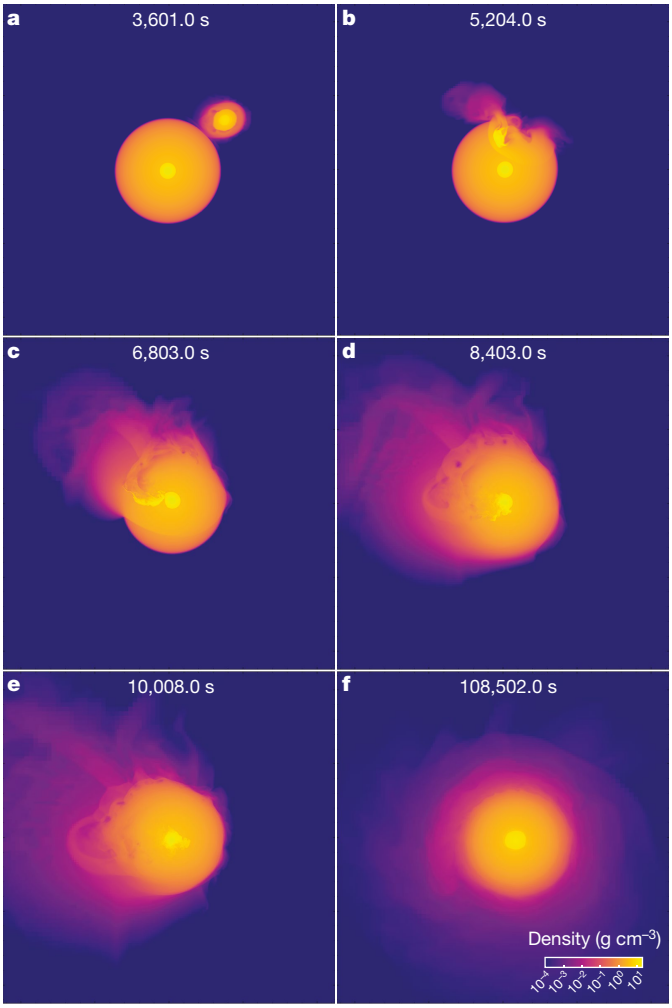


Fig. 3 | Two-dimensional snapshots of an off-centre collision between the proto-Jupiter with a $10M_{\oplus}$ rock/ice core and a $10M_{\oplus}$ impactor. **a**–**f**, Density contours in the orbital plane before the impact (**a**); during disruption and accretion of the impactor (**b**–**e**); at about 30 h after the impact (**f**). The time shown in each panel is measured since the start of the simulation. See Methods for detailed discussion.

A gradual accretion of planetesimals along with the runaway gas accretion may also produce a diluted core[15,28]. A relevant issue to be investigated elsewhere is whether the steep composition gradient needed to preserve the diluted core can also be established after a series of planetesimal-accretion events rather after than a single embryo's giant impact. Finally, extrasolar gas giant planets could also experience such giant impacts, which may explain the extremely large bulk metallicities of some giant exoplanets[29].

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1470-2.

Received: 29 September 2018; Accepted: 20 June 2019;
Published online 14 August 2019.

1. Bolton, S. J. et al. The Juno mission. *Space Sci. Rev.* **213**, 5–37 (2017).
2. Folkner, W. M. et al. Jupiter gravity field estimated from the first two Juno orbits. *Geophys. Res. Lett.* **44**, 4694–4700 (2017).
3. Wahl, S. M. et al. Comparing Jupiter interior structure models to Juno gravity measurements and the role of a dilute core. *Geophys. Res. Lett.* **44**, 4649–4659 (2017).
4. Debras, F. & Chabrier, G. New models of Jupiter in the context of Juno and Galileo. *Astrophys. J.* **872**, 100 (2019).
5. Pollack, J. B. et al. Formation of the giant planets by concurrent accretion of solids and gas. *Icarus* **124**, 62–85 (1996).
6. Ikoma, M., Nakazawa, K. & Emori, H. Formation of giant planets: dependences on core accretion rate and grain opacity. *Astrophys. J.* **537**, 1013–1025 (2000).
7. Helled, R. et al. Giant planet formation, evolution, and internal structure. *Protostars Planets* **VI**, 643 (2014).
8. Paardekooper, S.-J. & Mellema, G. Planets opening dust gaps in gas disks. *Astron. Astrophys.* **425**, L9–L12 (2004).
9. Levison, H. F., Thommes, E. & Duncan, M. J. Modeling the formation of giant planet cores. i. evaluating key processes. *Astron. J.* **139**, 1297–1314 (2010).
10. Bitsch, B. et al. Pebble-isolation mass: scaling law and implications for the formation of super-Earths and gas giants. *Astron. Astrophys.* **612**, A30 (2018).
11. Guillot, T., Stevenson, D. J., Hubbard, W. B. & Saumon, D. The interior of Jupiter. In *Jupiter: The Planet, Satellites and Magnetosphere* 35–57 (Cambridge Univ. Press, 2004).
12. Wilson, H. F. & Militzer, B. Solubility of water ice in metallic hydrogen: consequences for core erosion in gas giant planets. *Astrophys. J.* **745**, 54 (2012).
13. Stevenson, D. J. Structure of the giant planets: evidence for nucleated instabilities and post-formational accretion. *Lunar Planet. Sci. Conf.* **13**, 770–771 (1982).
14. Hori, Y. & Ikoma, M. Gas giant formation with small cores triggered by envelope pollution by icy planetesimals. *Mon. Not. R. Astron. Soc.* **416**, 1419–1429 (2011).
15. Lozovsky, M., Helled, R., Rosenberg, E. D. & Bodenheimer, P. Jupiter's formation and its primordial internal structure. *Astrophys. J.* **836**, 227 (2017).
16. Guillot, T. The interiors of giant planets: models and outstanding questions. *Annu. Rev. Earth Planet. Sci.* **33**, 493–530 (2005).
17. Nettelmann, N., Becker, A., Holst, B. & Redmer, R. Jupiter models with improved ab initio hydrogen equation of state (H-REOS.2). *Astrophys. J.* **750**, 52 (2012).
18. Helled, R. & Guillot, T. Interior models of Saturn: including the uncertainties in shape and rotation. *Astrophys. J.* **767**, 113 (2013).
19. Li, S.-L., Agnor, C. & Lin, D. N. C. Embryo impacts and gas giant mergers. I. Dichotomy of Jupiter and Saturn's core mass. *Astrophys. J.* **720**, 1161–1173 (2010).
20. Liu, S.-F., Agnor, C. B., Lin, D. N. C. & Li, S.-L. Embryo impacts and gas giant mergers—II. Diversity of hot Jupiters' internal structure. *Mon. Not. R. Astron. Soc.* **446**, 1685–1702 (2015).
21. Kokubo, E. & Ida, S. Oligarchic growth of protoplanets. *Icarus* **131**, 171–178 (1998).
22. Ida, S. & Lin, D. N. C. Toward a deterministic model of planetary formation. I. A desert in the mass and semimajor axis distributions of extrasolar planets. *Astrophys. J.* **604**, 388–413 (2004).
23. Zhou, J.-L. & Lin, D. N. C. Planetesimal accretion onto growing proto-gas giant planets. *Astrophys. J.* **666**, 447–465 (2007).
24. Ida, S., Lin, D. N. C. & Nagasawa, M. Toward a deterministic model of planetary formation. VII. Eccentricity distribution of gas giants. *Astrophys. J.* **775**, 42 (2013).
25. Fryxell, B. et al. FLASH: an adaptive mesh hydrodynamics code for modeling astrophysical thermonuclear flashes. *Astrophys. J. Suppl.* **131**, 273–334 (2000).
26. Berardo, D. & Cumming, A. Hot-start giant planets form with radiative interiors. *Astrophys. J.* **846**, L17 (2017).
27. Cumming, A., Helled, R. & Venturini, J. The primordial entropy of Jupiter. *Mon. Not. R. Astron. Soc.* **477**, 4817–4823 (2018).
28. Helled, R. & Stevenson, D. The fuzziness of giant planets' cores. *Astrophys. J.* **840**, L4 (2017).
29. Thorngren, D. P. & Fortney, J. J. Bayesian analysis of hot-Jupiter radius anomalies: evidence for ohmic dissipation? *Astron. J.* **155**, 214 (2018).

METHODS

A statistical *N*-body study of embryo collisions. We investigate the statistics of collisions between an emerging Jupiter and planetary embryos with the open-source *N*-body code REBOUND[30] version 3.6.2. To simulate the evolution of a planetary system we choose the built-in hybrid HERMES integrator, which uses the WHFast integrator[31] for the long-term dynamics and switches to the IAS15 integrator[32] when close encounters such as scattering and collisions happen (in recent updates of REBOUND, the HERMES integrator has been replaced by the MERCURIUS integrator, which offers a similar capability in a single scheme).

Our *N*-body simulations start from a coplanar configuration in which five $10M_\oplus$ planetary embryos ($M_\mathrm{p} = 10M_\oplus$) orbit the Sun ($M_* = 1M_\odot \approx 3.3 \times 10^5 M_\oplus$) on circular prograde orbits. The embryo at 5.2 AU from the Sun grows into a Jupiter-mass planet at the end of the simulation. Initially, two embryos are placed interior to Jupiter's orbit and the other two embryos are placed exterior to Jupiter's orbit. The orbital separation between any two adjacent embryos i and $i+1$ is determined by a dimensionless number:

$$k = \frac{2a_i}{r_\mathrm{H}}\left(\frac{a_{i+1}-a_i}{a_{i+1}+a_i}\right) \tag{1}$$

where a_i and a_{i+1} are the semi-major axes of each embryo, and $r_\mathrm{H} = a_i\left(\frac{M_\mathrm{p}}{3M_*}\right)^{1/3}$ is the Hill radius of embryo i. It is convenient to express equation (1) in terms of $q = a_{i+1}/a_i$, the ratio of semi-major axis between embryos i and $i+1$:

$$k = 2\left(\frac{3}{\mu}\right)^{1/3}\left(\frac{q-1}{q+1}\right) \tag{2}$$

where $\mu = M_\mathrm{p}/M_* \approx 3 \times 10^{-5}$ is the mass ratio between the embryo and the Sun. A larger k will give rise to a wider separation, that is, a more dynamically stable configuration. Extended Data Table 1 summarizes the locations of all embryos for a given parameter k in our *N*-body simulation suite. In addition, we also consider a configuration in which all four embryos are beyond Jupiter's orbit.

At the onset of the simulation, the runaway gas accretion of Jupiter's core starts. The mass accretion rate is an exponential decay function characterized by an exponential time parameter t_grow ranging from 0.1 million years to 0.5 million years in this study. At a given time t, the mass of an emerging Jupiter is determined by

$$M(t) = M_\mathrm{J} - (M_\mathrm{J} - M_\mathrm{p}) \times \mathrm{e}^{-t/t_\mathrm{grow}} \tag{3}$$

where $M_\mathrm{J} = 317.8M_\oplus$ is one Jupiter mass. In this model, Jupiter quickly acquires more than 90% of its mass within $3t_\mathrm{grow}$ and steadily gains another a few per cent of its mass until $t = 10t_\mathrm{grow}$. For simplicity, we assume that all of the other four embryos do not grow, since the typical hydrostatic growth stage of an embryo before it enters the runaway gas accretion phase is a few million years long, during which the embryo mass hardly increases.

Size is another crucial factor because a larger planetary cross-section can boost the probability of collisions. We adopt the mean density of the Earth for embryos, so their sizes are $R_\mathrm{p} \approx 2.15R_\oplus$, where $R_\oplus$ is Earth's mean radius. For the emerging Jupiter, its mean density could be as low as half of its present-day value. We use the parameter f to describe the degree of inflation.

Thus, we design a simple classification for our *N*-body simulation suite with three free parameters k, t_grow and f. For each combination set of (k, t_grow, f), we run thousands of simulations with other orbital parameters (such as true anomaly or argument of periapsis) randomly chosen between 0 and 2π.

At the end of an *N*-body simulation ($t = 10t_\mathrm{grow}$), a planetary embryo may remain bound to the Sun with considerable changes in its orbit, or coalesce with Jupiter and other embryos, or escape from the system after a close encounter. The statistics of the final outcomes of four planetary embryos under the influence of an emerging Jupiter is shown in Extended Data Fig. 1. The results are grouped by different parameters to compare their impacts. In all subsets of our *N*-body simulations, we observe an efficient pathway towards planetary embryos colliding with an emerging Jupiter.

Because embryos are equally distributed on both sides of Jupiter's orbit (except for the last group that starts with all embryos in the 'Outward' state), the results suggest that embryos both interior or exterior to Jupiter could collide with Jupiter within the simulation time. However, embryos beyond Jupiter may have a slightly larger chance of striking Jupiter given that there are fewer embryos remaining in the 'Outward' state at the end of the simulation. Of the three key parameters, orbital tightness characterized by k has the most substantial role in affecting the collision probability. For the same orbital configuration, Jupiter inflation factor f can slightly change the collision rate. However, Jupiter's accretion history, determined by t_grow, has the least influence on the results.

We also analyse the distribution of collision angle using our *N*-body simulation suite. The histograms of collision angles are plotted in Extended Data Fig. 2. Each histogram represents a detailed breakdown of 'Merge' events of a simulation set presented in Extended Data Fig. 1. Unlike collisions between similar-sized planetary bodies, in which 45° collisions are common[33], the statistical results suggest that half of the merger events have collision angles less than about 30° in all cases we investigated. We suggest that low-angle impacts are very frequent because of Jupiter's strong gravitational focusing effect.

It is often useful to define a two-body escape velocity as

$$V_\mathrm{esc} = \left(\frac{2G(M_\mathrm{J}+M_\mathrm{p})}{R_\mathrm{R}+R_\mathrm{p}}\right)^{1/2} \tag{4}$$

which is around 51 km s^{-1} for the proto-Jupiter and the $10M_\oplus$ impactor studied in the hydrodynamic simulation. In general, an embryo's impact velocity V_imp is related to V_esc as well as to the local Keplerian velocity V_Kepler. Gravitational perturbation during close encounters can produce an impact velocity with a magnitude up to the escape velocity[34]. On the other hand, the Keplerian orbital velocity gives rise to the random velocity dispersion during impacts. At Jupiter's current location, $V_\mathrm{Kepler} \approx 13$ km s^{-1} is much smaller than V_esc, so the impact velocity V_imp is approximately at the escape velocity V_esc. Indeed, we find the impact velocity is quantitatively similar to V_esc rather than V_Kepler, although V_imp is always slightly smaller than V_esc in the *N*-body simulation suite, because initial separations between Jupiter and embryos are finite (a two-body system has a negative gravitational potential energy).

This simple statistical model may be improved to enable comparison with other formation models of the outer Solar system in future studies. For example, because Jupiter's inward migration is much slower than those planetary embryos, the presence of Jupiter in the Solar nebula acts like a barrier for inward-migrating planetary embryos formed exterior to Jupiter[35]. Consequently, collisions among those planetary embryos may become frequent and some of those events may eventually form Uranus and Neptune[36].

Hydrodynamic simulations. Our three-dimensional hydrodynamic simulation of giant impacts between a proto-Jupiter and a planetary embryo is based on the framework of the Eulerian FLASH code[25], which utilizes the adaptive-mesh refinement. The setup of giant impact simulations has been described in our previous study[37]. Here we briefly describe the model of the planetary interior. Both the proto-Jupiter and the impactor are modelled with a three-layer structure: a silicate inner core, an icy outer core, and a H-He envelope. We calculate two thermodynamic properties (density and internal energy) of silicate and ice material and their velocities using the governing continuity, momentum and energy equation. For computational efficiency, these quantities are converted into pressure and temperature with the Tillotson EOS[38]. The mass fraction between ice to silicate is assumed to be 2.7 according to that of the proto-Sun (2–3). In addition, the H-He EOS is modelled with an $n = 1$, $\gamma = 2$ polytropic relation, where n and γ are the polytropic and adiabatic indexes. Although this idealized treatment ignores effects such as the H-He phase transition and separation, it matches the density profile of Jupiter's envelope calculated with ab initio EOS[39] reasonably well and is good enough for dynamic processes that happen within a few hours (see detailed discussion below).

Collisions between a proto-Jupiter with a $10M_\oplus$ core and a $10M_\oplus$ embryo. From *N*-body simulations we learn that most collisions have collision angles of less than 30°, so we first study the head-on collision as one of the representative cases in the main text and the consequence is shown in Fig. 1. We also plot its two-dimensional counterpart in Extended Data Fig. 3. The general behaviour of head-on collisions has been studied extensively in previous work[20,37]. To recapitulate, the heavy material of the impactor can penetrate Jupiter's gaseous envelope and smash into its core as a whole. As a result, Jupiter's core gets completely destroyed after the impact. The release of a large amount of energy inside the proto-Jupiter drives largescale turbulence and the primordial compact core is subsequently eroded. We compare the enclosed internal energy of Jupiter as a function of radius before and after the impact. The results are shown in Extended Data Fig. 4. Although Jupiter gains internal energy through the release of kinetic and gravitational energy of the impactor as well as the impactor's own internal energy, the core region is hardly heated. In fact, there is even a small decrease in internal energy inside the core region immediately after the impact, possibly due to mixing with H-He. The analysis suggests that the impactor dumps most of its energy outside the original core region.

Our simplified EOS for H-He causes less efficient dissipation of the impactor's kinetic energy within the H-He envelope. As vigorous mixing between H-He and core material, however, is driven by a merger between the core of a proto-Jupiter and an impactor, we can expect the formation of a diluted core to occur regardless of EOS models. In addition, a temperature profile inside a core is not strongly

affected by the choice of a H-He EOS model because the impact causes only a small change in internal energy inside the core.

To illustrate the effects of off-centre collisions, we run the same setup of simulation except that the collision angle is at 45°. The consequence is shown in Fig. 3. Because the initial impact velocity is close to the escape velocity, the impactor misses Jupiter's core and overshoots until Jupiter's gravitational force pulls it back. During its course, the impactor gradually loses angular momentum and gets torn apart. The remnant is gently accreted by Jupiter's rock/ice core later on. As a result, the impact has little influence on Jupiter's core–envelope structure.

A head-on collision between a proto-Jupiter with a massive core and a small impactor. In addition, we perform a head-on collision between a proto-Jupiter with a massive primordial core of $17M_{\oplus}$ and a $1M_{\oplus}$ impactor, which is composed of pure silicate, at the same impact velocity. The total amount of heavy elements is the same as that in previous head-on and off-centre models (hereafter, case 1 and case 2). Unlike case 1, the impactor disintegrates in the proto-Jupiter's envelope before making contact with the core. A strong shockwave induced by the entry of the impactor propagates throughout the entire planet and deforms the core (see Extended Data Fig. 5c). After the impact, a small fraction of H-He (only about 5 wt%) is mixed into the proto-Jupiter's core owing to a weak impact-induced oscillation. As a result, the central density of the core still decreases by a factor of two-thirds. Although the core–envelope boundary spreads out slightly, a steep density gradient between the core and the H-He envelope is preserved, leading to the retention of a compact, massive core.

To summarize, only in case 1 we observe a smooth transition between the core and the H-He envelope after the impact, because the impactor is massive and hits Jupiter's core directly. However, in both case 2 and case 3, because the impactor is unable to collide with the core as an integrated body, the proto-Jupiter's core becomes slightly enriched in H-He after it gets restored from deformation. Therefore, we conclude that neither a small impactor nor an off-centre collision is able to form a large diluted core. A proto-Jupiter with a primordial compact core must have experienced a catastrophic nearly head-on collision with a large embryo if its present-day Jupiter has a massive, diluted core. A more comprehensive parameter study, including a range of impactor mass and speed as well as off-centre collisions, will be presented elsewhere.

Post-impact thermal evolution. We simulate Jupiter's long-term evolution after the giant impact in order to identify the evolutionary paths that lead to a diluted core structure at present-day. The planetary evolution is modelled using the one-dimensional stellar evolution code Modules for Experiments in Stellar Astrophysics (MESA), where the planet is assumed to be spherically symmetric and in hydrostatic equilibrium[40–43]. The evolution is modelled with a modification to the EOS (S.M., A. Cumming & R.H.; manuscript in preparation), where the H-He EOS is based on SCVH[44] with an extension to lower pressures and temperatures, and the heavy-element (H_2O/SiO_2) EOS is QEOS[45,46]. Conductive opacities are from ref. [47], and the molecular opacity is from ref. [48].

The planetary evolution is governed by the energy transport in the interior, which can occur via radiation, conduction or convection. We use the standard Ledoux criterion[49] to determine whether a region with composition gradients is stable against convection, that is, $\nabla_T < \nabla_{\rm adiab} + B$, where $\nabla_T = {\rm d}(\log T)/{\rm d}(\log P)$, with $\nabla_{\rm adiab}$ and B being the adiabatic temperature and composition gradient, respectively. If the composition gradient is such that the mean molecular weight increases towards the planetary centre, then $B > 0$ and the composition gradient could inhibit convection. For a homogeneous planet, $B = 0$ and the Ledoux criterion reduces to the Schwarzschild criterion $\nabla_T < \nabla_{\rm adiab}$. A region that is Ledoux stable but Schwarzschild unstable could develop semi-convection. In that case, double-diffusive processes can lead to additional mixing[50].

In the planet evolution code, convective mixing is treated via the mixing length theory (MLT), which provides a recipe to calculate ∇_T and the diffusion coefficient, fully determining the convective flux. The MLT requires the knowledge of a mixing length $l_{\rm mix} = \alpha_{\rm MLT} H_P$, where H_P is the pressure scale height and $\alpha_{\rm MLT}$ is a dimensionless parameter. The expected value of $\alpha_{\rm MLT}$ for planets is poorly constrained. Following previous work on Jupiter's evolution with convective mixing[51] we use $\alpha_{\rm MLT} = 0.1$ as our baseline. We find that the mixing is relatively insensitive to the choice of the mixing length within about an order of magnitude. This is because its value does not directly determine when mixing occurs, but rather the mixing efficiency. To investigate the sensitivity of the results on this parameter we also included a model with $\alpha_{\rm MLT} = 10^{-3}$. Although our conclusions on the diluted core are robust, a detailed and rigorous investigation into mixing in giant planets is clearly desirable, and will be presented in future work (S.M., A. Cumming & R.H.; manuscript in preparation).

The case of semi-convection is treated as a diffusive process[52], which requires the calculation of the temperature gradient and diffusion coefficient in the semi-convective region. The recipe includes a free parameter that can be interpreted as the layer-height of the double-diffusive region[53,54], which is unknown and could range over a few orders of magnitude. In the case where we include semi-convection, we set the value to 10^{-5} pressure scale heights, which is an value intermediate in the range given in the literature[55].

The hydro-simulation of the giant impact sets the post-impact composition profile to be used by the evolution model. The initial temperature profile is crucial for determining the energy transport for the subsequent evolution. Because the proto-Jupiter's thermal state at the time of impact is unknown, we consider various initial temperature profiles and explore how the evolution is affected by this choice. Giant planet formation calculations estimate the central temperature of the proto-Jupiter to be around 10^4 K (ref. [27]). The exact temperature, however, is unknown and can change by tens of per cent (a factor of a few). For determining the convective mixing efficiency such factors can lead to large differences in the long-term evolution and the final internal structure. Also, recent work has shown that accounting for the accretion shock during the runaway gas accretion phase can lead to a radiative envelope and a non-monotonic temperature profile in the deep interior[26,27]. We include this possibility in one of our models (model H-radenv). Our nominal models use $\alpha_{\rm MLT} = 0.1$, with no semi-convection, and the heavy elements represented by water. A summary of the model parameters is given in Extended Data Table 2.

In Extended Data Fig. 6 we present the starting models that are evolved to Jupiter's present age. The solid and dashed lines correspond to the head-on and oblique (at an angle of 45°) collisions, respectively. The temperatures are increasing towards the interior for all models except H-radenv, as explained above. A temperature inversion occurs in the deep interior, corresponding to the location of the accretion shock during early runaway gas accretion. We note that in this model, the temperature inversion occurs around the same region as the composition gradient, providing additional support against convection. Although the exact location of the temperature jump is unknown, it is expected to be relatively narrow. It is limited by the so-called crossover mass, which a giant planet must reach in order to enter the runaway gas accretion phase[56]. As the heavy-element fraction increases, the interior becomes hotter as a result of the change in opacity and the increase in density. If the collision is head-on, the composition gradient is shallower and extends farther into the envelope.

Extended Data Figs. 7, 8 show the density profiles after 4.56 Gyr of evolution for the head-on and oblique collisions, respectively. The crucial influence of the initial thermal profile on the mixing is clear: For the $\log[T_{\rm central}\,({\rm K})] = 4.7$, where $T_{\rm central}$ is the temperature of the centre of Jupiter, head-on collision case (model H-4.7), the end result is a fully homogeneous Jupiter without a core. For the oblique impact, even the very steep composition gradient, with the highest temperatures, is insufficient to inhibit substantial mixing of the deep interior. The intermediate temperature profiles lead to varying degrees of mixing. In general, the head-on collision results in an extended core that is highly enriched in H-He, while for the oblique impact the core is more compact and less diluted. Despite a substantial fraction of the proto-Jupiter being very hot in the model H-radenv, there is not enough mixing to erase the composition gradient. In this case, the envelope is radiative at early times when mixing would be most efficient. If a lower mixing length is chosen (model H-4.5-lowα), the composition gradient is less eroded and extends farther into the envelope. Because the energy transport is also affected by the chosen mixing length, Jupiter's interior is hotter and less dense compared to that in model H-4.5.

Model H-4.5-semiconv is the same as model H-4.5 but allowing semi-convective mixing with a layer height of 10^{-5} pressure scale heights. In this case, semi-convection is insufficient to overcome the stabilizing composition gradient. Although some additional mixing occurs, particularly at early times, there are no semi-convective regions towards the end of the evolution. In other words, the final interior structure is such that the radiative regions are both Schwarzschild and Ledoux stable. This demonstrates that when semi-convection is included we can also infer a Jupiter with a diluted core.

To completely erase the composition gradient created by the giant impact, the impact must be head-on and the post-impact interior needs to be very hot (about 50,000 K) with the heavy elements represented by water (model H-4.7). In all the other models we consider, the stabilizing effect of the post-impact heavy-element distribution is inhibiting the development of convective instabilities, resulting in an inhomogeneous Jupiter. Therefore, the typical outcome of the calculation is an interior structure that is not fully mixed and is characterized by several radiative–convective interfaces. Interestingly, the development of these interfaces seems to be a frequent occurrence when modelling Jupiter's evolution with composition gradients[51] (S.M., A. Cumming & R.H.; manuscript in preparation). If the core is defined as the region that is substantially richer in heavy elements than the envelope, then most of our models imply that Jupiter has a diluted core extending to about 30%—50% of the planet's radius. All of the oblique collisions lead to a relatively compact core since the initial composition gradient is very steep.

Figure 2b shows the models that best match the diluted-core density profile from ref. [3] (models H-4.5-rock, H-4.5 and H-radenv). We find that for the head-on collision, a post-impact central temperature of about 30,000 K leads to a current-state

Jupiter with a diluted core (models H-4.5 and H-4.5-rock). If the heavy elements are represented by rock (SiO_2), the diluted core extends farther into the envelope and is thus more consistent with the Jupiter structure presented in ref. [3]. Another pathway to the diluted core is when Jupiter's deep interior is radiative, owing to the accretion shock, as predicted by recent giant planet formation models[27] (model H-radenv). Videos that demonstrate the planetary evolution for three selected cases can be found in the Supplementary Information.

Data availability

The datasets generated and analysed during the current study are available from the corresponding authors upon reasonable request.

Code availability

The FLASH code is publicly available for download at http://flash.uchicago.edu/site/flashcode. The implementation of giant impact simulations in the framework of FLASH is available upon request. The REBOUND code is publicly available for download at https://github.com/hannorein/rebound. The MESA code is an open source stellar evolution code and is publicly available at http://mesa.sourceforge.net. The modified version of the MESA code is not yet ready for public release—it will be presented in future work (S.M., A. Cumming & R.H.; manuscript in preparation). Gnuplot, Jupyter Notebook, Mathematica, VisIt and yt python packages were used for data reduction and presentation in this study.

30. Rein, H. & Liu, S.-F. REBOUND: an open-source multi-purpose *N*-body code for collisional dynamics. *Astron. Astrophys.* **537**, A128 (2012).
31. Rein, H. & Tamayo, D. WHFAST: a fast and unbiased implementation of a symplectic Wisdom–Holman integrator for long-term gravitational simulations. *Mon. Not. R. Astron. Soc.* **452**, 376–388 (2015).
32. Rein, H. & Spiegel, D. S. IAS15: a fast, adaptive, high-order integrator for gravitational dynamics, accurate to machine precision over a billion orbits. *Mon. Not. R. Astron. Soc.* **446**, 1424–1437 (2015).
33. Asphaug, E., Agnor, C. B. & Williams, Q. Hit-and-run planetary collisions. *Nature* **439**, 155–160 (2006).
34. Lin, D. N. C. & Ida, S. On the origin of massive eccentric planets. *Astrophys. J.* **477**, 781–791 (1997).
35. Izidoro, A., Raymond, S. N., Morbidelli, A., Hersant, F. & Pierens, A. Gas giant planets as dynamical barriers to inward-migrating super-Earths. *Astrophys. J.* **800**, L22 (2015).
36. Izidoro, A., Morbidelli, A., Raymond, S. N., Hersant, F. & Pierens, A. Accretion of Uranus and Neptune from inward-migrating planetary embryos blocked by Jupiter and Saturn. *Astron. Astrophys.* **582**, A99 (2015).
37. Liu, S.-F., Hori, Y., Lin, D. N. C. & Asphaug, E. Giant impact: an efficient mechanism for the devolatilization of super-Earths. *Astrophys. J.* **812**, 164 (2015).
38. Melosh, H. J. *Impact Cratering: A Geologic Process* (Oxford Univ. Press, 1989).
39. Liu, S.-F., Guillochon, J., Lin, D. N. C. & Ramirez-Ruiz, E. On the survivability and metamorphism of tidally disrupted giant planets: the role of dense cores. *Astrophys. J.* **762**, 37 (2013).
40. Paxton, B. et al. Modules for Experiments in Stellar Astrophysics (MESA). *Astrophys. J. Suppl.* **192**, 1–110 (2011).
41. Paxton, B. et al. Modules for experiments in stellar astrophysics (MESA): planets, oscillations, rotation, and massive stars. *Astrophys. J. Suppl.* **208**, 4 (2013).
42. Paxton, B. et al. Modules for Experiments in Stellar Astrophysics (MESA): binaries, pulsations, and explosions. *Astrophys. J. Suppl.* **220**, 1–43 (2015).
43. Paxton, B. et al. Modules for Experiments in Stellar Astrophysics (MESA): convective boundaries, element diffusion, and massive star explosions. *Astrophys. J. Suppl.* **234**, 34 (2018).
44. Saumon, D., Chabrier, G. & van Horn, H. M. An equation of state for low-mass stars and giant planets. *Astrophys. J. Suppl.* **99**, 713 (1995).
45. More, R. M., Warren, K. H., Young, D. A. & Zimmerman, G. B. A new quotidian equation of state (QEOS) for hot dense matter. *Phys. Fluids* **31**, 3059–3078 (1988).
46. Vazan, A., Kovetz, A., Podolak, M. & Helled, R. The effect of composition on the evolution of giant and intermediate-mass planets. *Mon. Not. R. Astron. Soc.* **434**, 3283–3292 (2013).
47. Cassisi, S., Potekhin, A. Y., Pietrinferni, A., Catelan, M. & Salaris, M. Updated electron-conduction opacities: the impact on low-mass stellar models. *Astrophys. J.* **661**, 1094–1104 (2007).
48. Freedman, R. S., Marley, M. S. & Lodders, K. Line and mean opacities for ultracool dwarfs and extrasolar planets. *Astrophys. J. Suppl.* **174**, 504–513 (2008).
49. Ledoux, W. P. Stellar models with convection and with discontinuity of the mean molecular weight. *Astrophys. J.* **105**, 305 (1947).
50. Rosenblum, E., Garaud, P., Traxler, A. & Stellmach, S. Turbulent mixing and layer formation in double-diffusive convection: three-dimensional numerical simulations and theory. *Astrophys. J.* **731**, 66 (2011).
51. Vazan, A., Helled, R. & Guillot, T. Jupiter's evolution with primordial composition gradients. *Astron. Astrophys.* **610**, L14 (2018).
52. Langer, N., Sugimoto, D. & Fricke, K. J. Semiconvective diffusion and energy transport. *Astron. Astrophys.* **126**, 207–208 (1983).
53. Wood, T. S., Garaud, P. & Stellmach, S. A new model for mixing by double-diffusive convection (semi-convection). II. The transport of heat and composition through layers. *Astrophys. J.* **768**, 157 (2013).
54. Radko, T. et al. Double-diffusive recipes. Part I: Large-scale dynamics of thermohaline staircases. *J. Phys. Oceanogr.* **44**, 1269–1284 (2014).
55. Leconte, J. & Chabrier, G. A new vision of giant planet interiors: impact of double diffusive convection. *Astron. Astrophys.* **540**, A20 (2012).
56. Baraffe, I., Chabrier, G., Fortney, J. J. & Sotin, C. Planetary internal structures. *Protostars Planets* **VI**, 763–786 (2014).

Acknowledgements We thank S. M. Wahl and Y. Miguel for sharing their results with us. We thank A. Cumming for technical support and discussions. We thank J. J. Fortney, P. Garaud and H. Rein for conversations. S.-F.L. acknowledges the support and hospitality provided by the Aspen Center for Physics during the early stage of this work. D.L. thanks the Institute for Advanced Study, Princeton, the Institute of Astronomy and Department of Applied Mathematics and Theoretical Physics, Cambridge University, for support and hospitality while this work was being completed. R.H. acknowledges support from SNSF grant number 200021_169054. A.I. acknowledges support from the National Aeronautics and Space Administration under award number 80NSSC18K0828 and from the National Science Foundation under grant number AST-1715719.

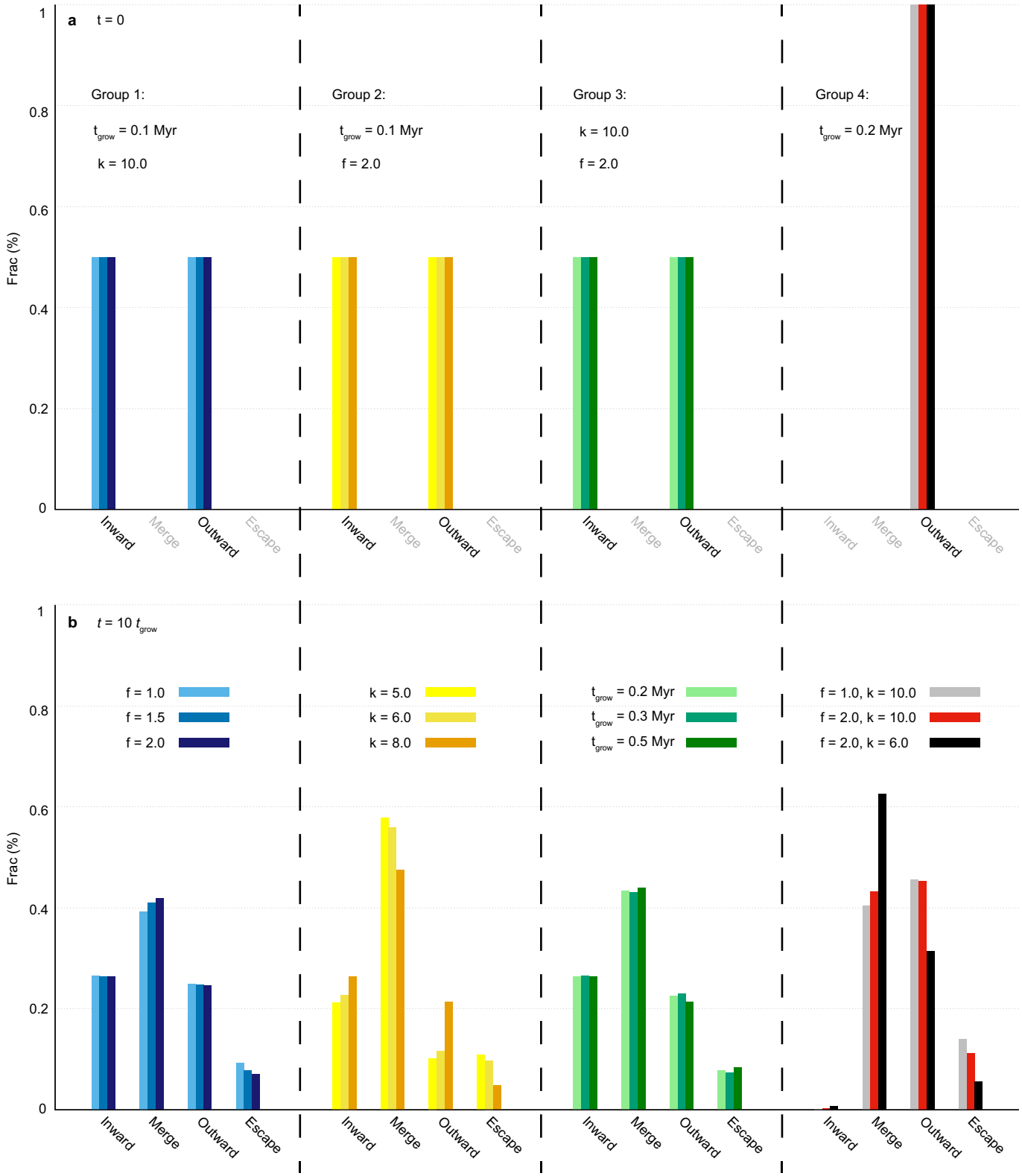
Author contributions D.L. had the idea of the impact scenario. S.-F.L. and A.I. examined its feasibility. S.-F.L. coordinated this study. S.-F.L. and Y.H. designed and analysed the hydrodynamic simulations. X.Z. and S.-F.L. performed and analysed the *N*-body simulations. S.M. and R.H. designed the long-term thermal evolution study. All authors contributed to discussions, as well as to editing and revising the manuscript.

Competing interests The authors declare no competing interests.

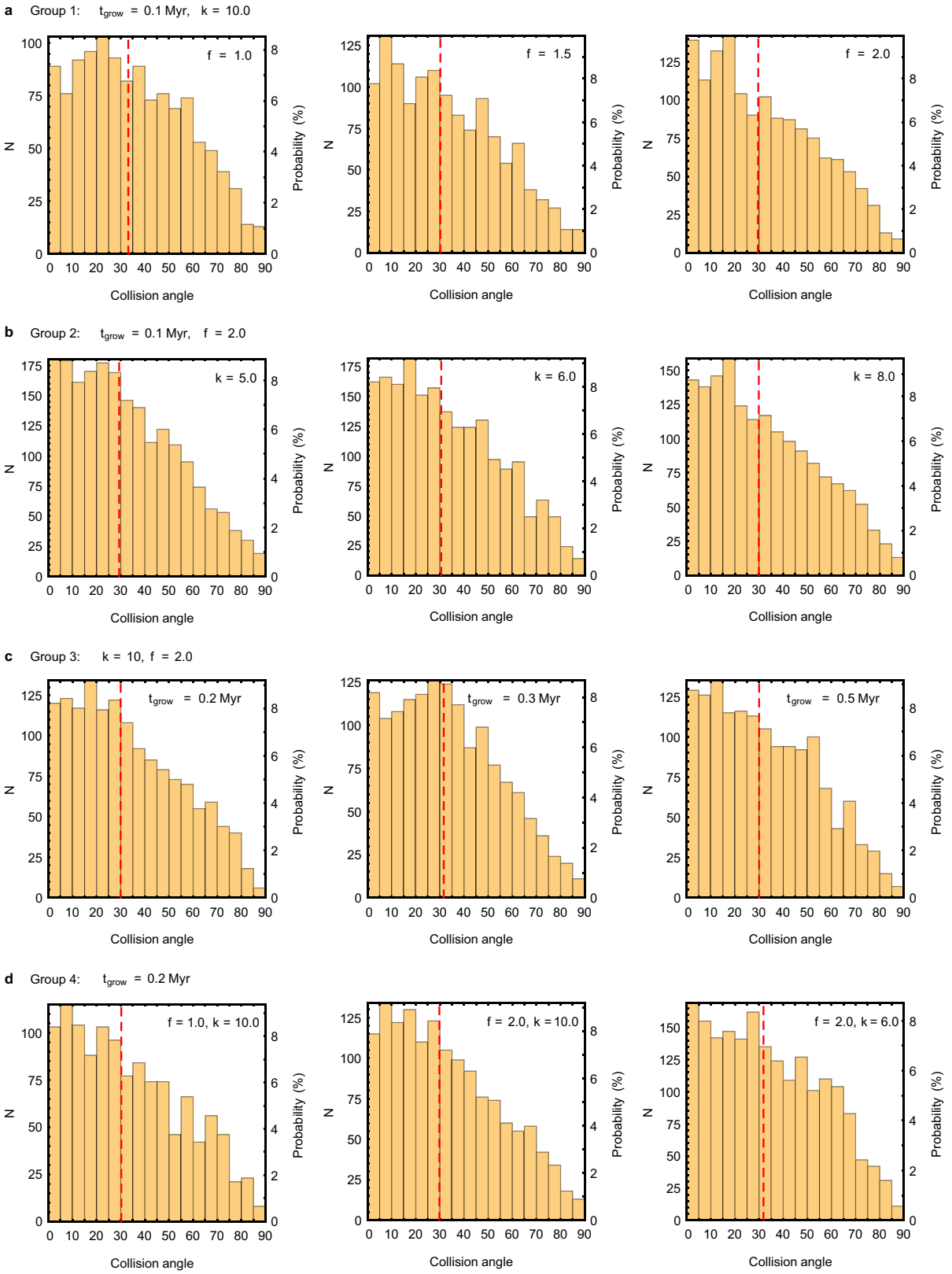
Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1470-2.

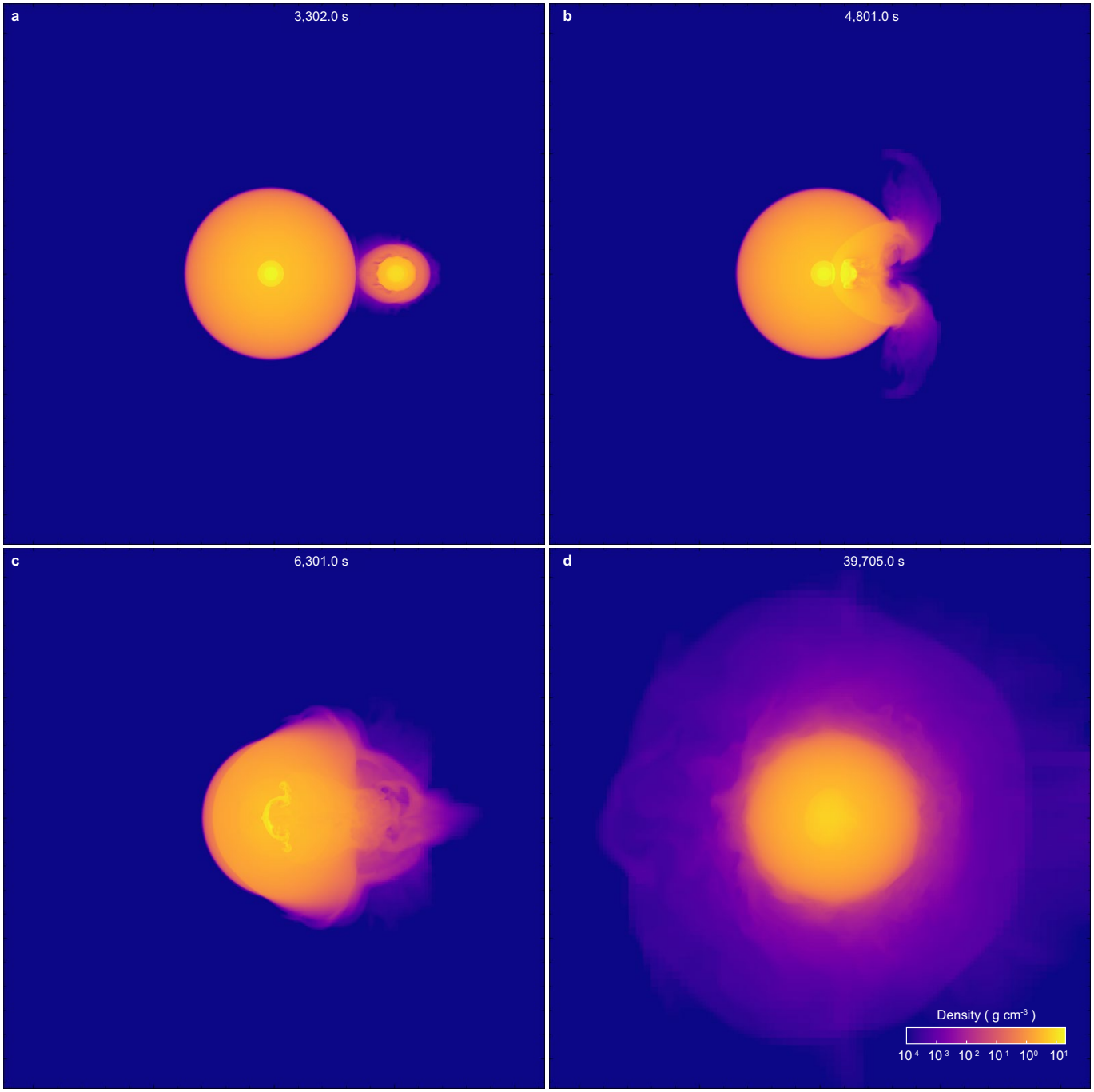
Correspondence and requests for materials should be addressed to S.-F.L.



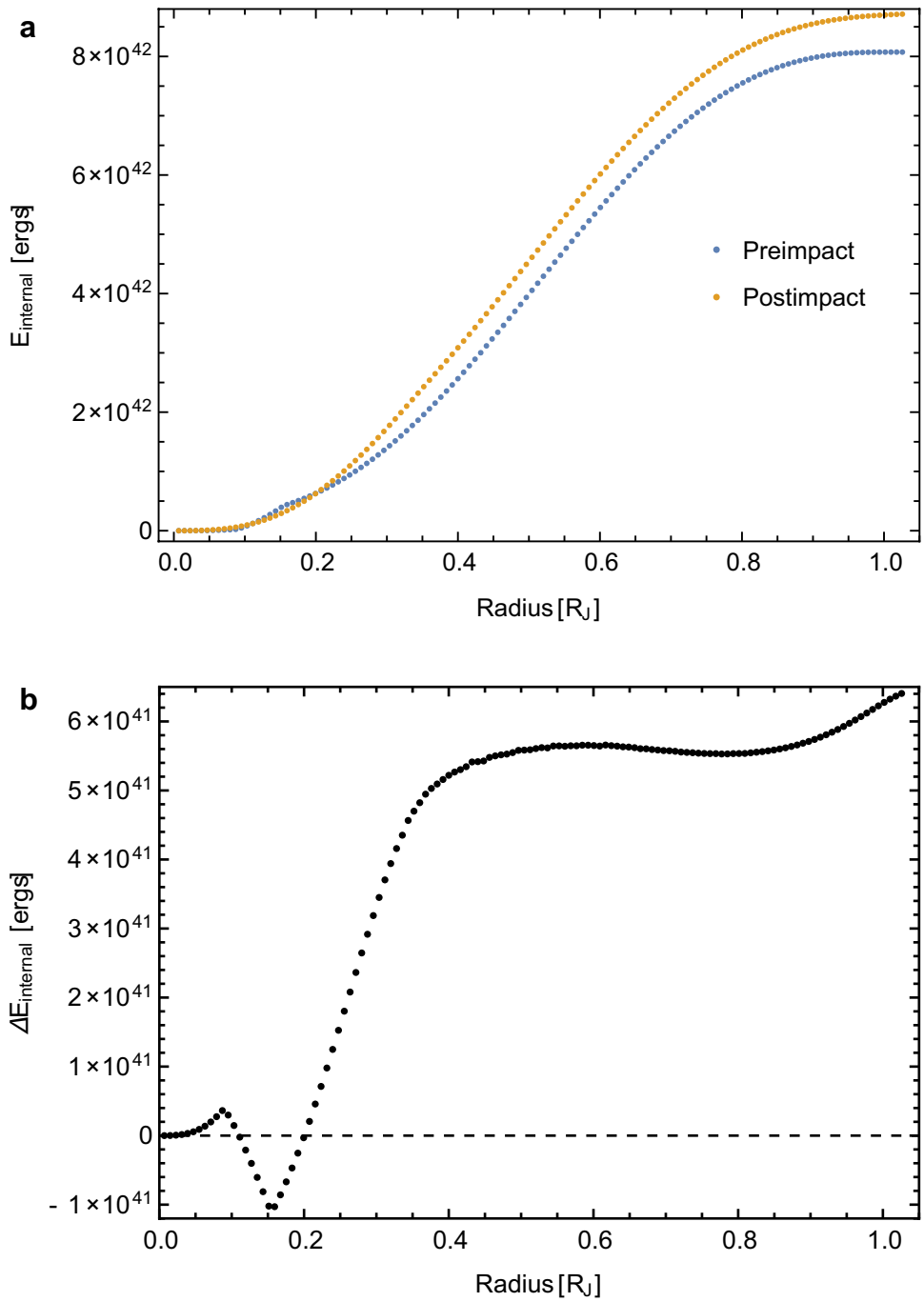
Extended Data Fig. 1 | Statistics of outcomes of four planetary embryos under the influence of an emerging Jupiter. a, The initial configurations of four planetary embryos divided into four groups based on the fixed parameters shown under the group numbers. In groups 1–3, half of the embryos are placed inside Jupiter's orbit (labelled 'Inward'); the other half are outside Jupiter's orbit (labelled 'Outward'). In group 4, all embryos are outside Jupiter's orbit. The exact location of every embryo is shown in Extended Data Table 1. **b**, The statistical outcomes of the dynamic evolution after $10t_{grow}$. Jupiter's growth can substantially modify the orbits of those embryos. Some embryos collided with Jupiter (labelled 'Merge'), and some have been ejected from the Solar System (labelled 'Escape'). Colours indicate different choices of the free parameters (inflation factor f and orbital separation factor k; see methods section 'A statistical N-body study of embryo collisions') as shown for each group. The height of each bar ('Frac') indicates the percentage of each state.



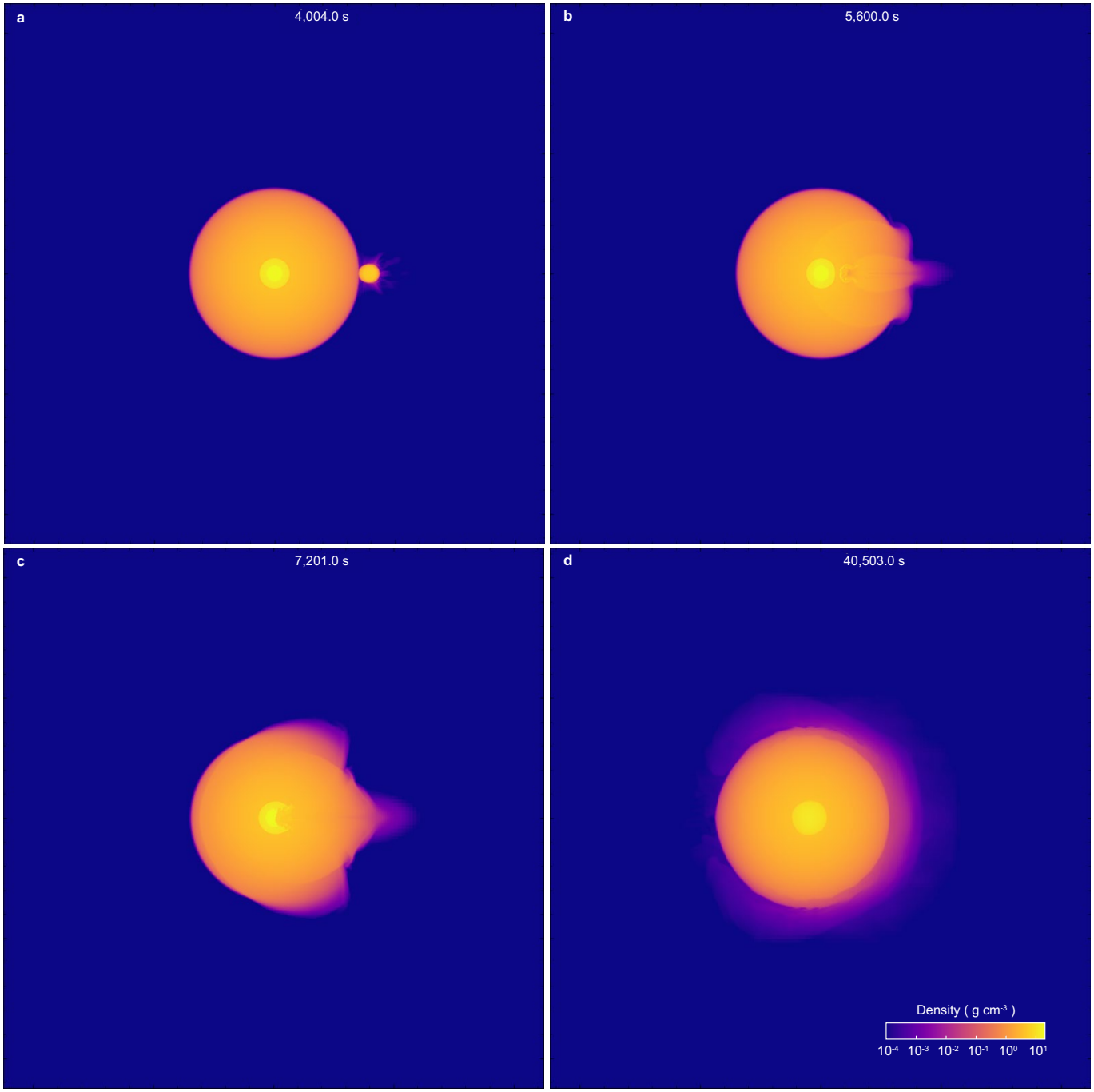
Extended Data Fig. 2 | Histograms of collision angles of each dataset presented in Extended Data Fig. 1. a, Group 1; **b**, group 2; **c**, group 3; **d**, group 4. The bin size is 5°, and there are 18 bins in each plot. The collision angle is measured in degrees. The red dashed line indicates the median value in each case. The results suggest that head-on collisions are more common (greater percentage probability) than grazing collisions. Each case has a different *N*, but they all fall in the range between 1,000 to 1,500.



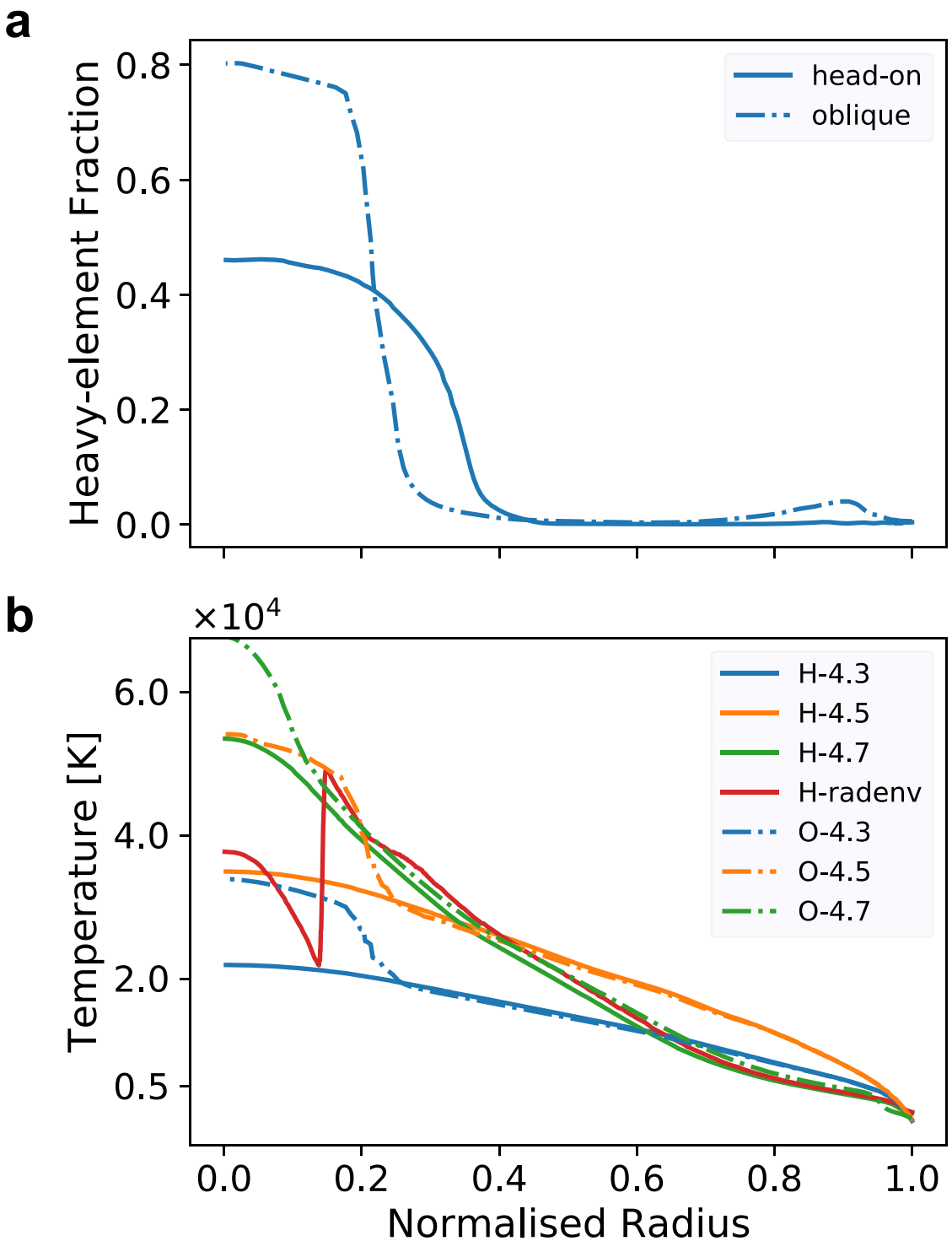
Extended Data Fig. 3 | Two-dimensional snapshots of a merger between the proto-Jupiter with a $10M_{\oplus}$ rock/ice core and a $10M_{\oplus}$ impactor. **a**, Density contours in the orbital plane before the impact; **b**, before the impactor arriving at the core; **c**, after the destruction of the core; **d**, at about 10 h after the impact. Time in each panel is measured since the start of the simulation.



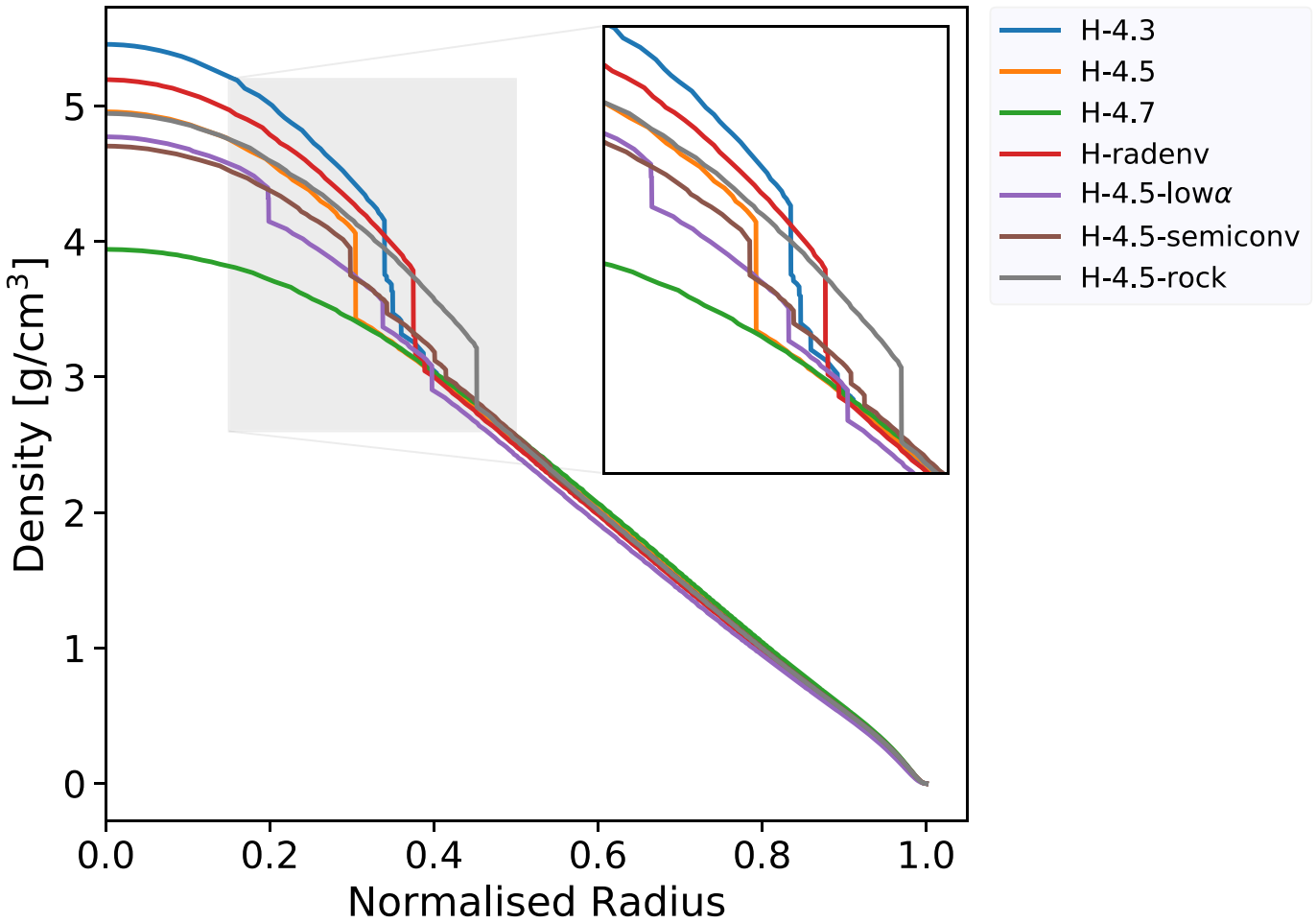
Extended Data Fig. 4 | The change of internal energy caused by the merger between the proto-Jupiter with a $10M_{\oplus}$ rock/ice core and a $10M_{\oplus}$ impactor. a, The enclosed internal energy $E_{\rm internal}$ of Jupiter before and after the impact as a function of radius. **b**, The net change of enclosed internal energy $\Delta E_{\rm internal}$ of Jupiter as a function of radius.



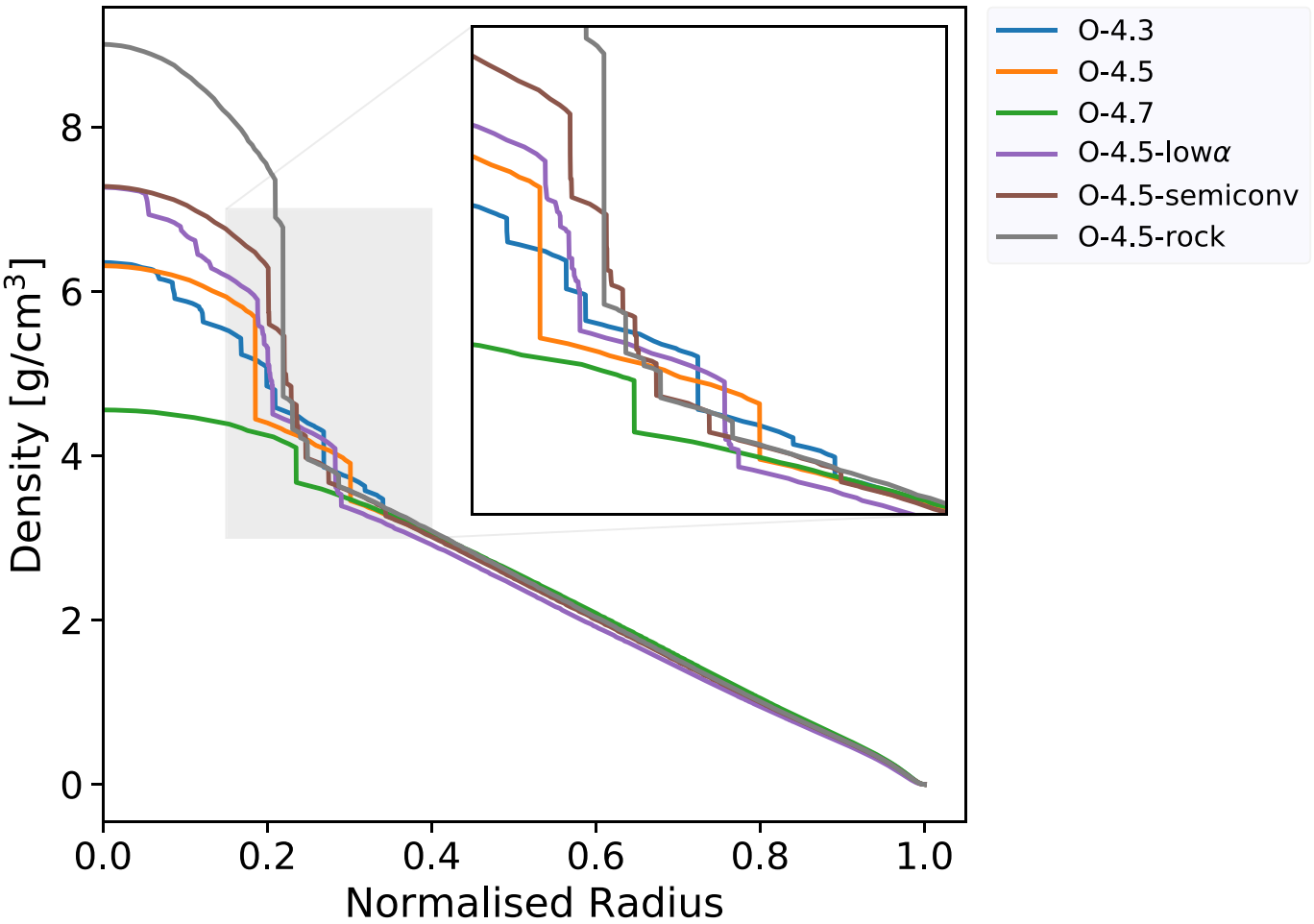
Extended Data Fig. 5 | Two-dimensional snapshots of a merger between the proto-Jupiter with a $17M_{\oplus}$ core and a $1M_{\oplus}$ impactor. a, Density contours in the orbital plane before the impact; **b**, before the impactor arriving at the core; **c**, after the merger with the core; **d**, at about 10 h after the impact. Time in each panel is measured since the start of the simulation.



Extended Data Fig. 6 | Initial conditions for post-impact evolution. **a**, The initial post-impact heavy-element profile; **b**, temperature profiles of the models used for the thermal evolution. The heavy-element distribution is taken from the hydro-simulation 10 h after the giant impact. Solid lines correspond to a head-on collision, while dashed-dotted lines show the result of an oblique collision at a 45° angle. The colours depict models with different initial thermal states. See text and Extended Data Table 2 for further details. The radius is normalized by the present-day radius of Jupiter R_J.



Extended Data Fig. 7 | Density versus normalized radius for the head-on collision after 4.56 Gyr of evolution. The colours correspond to distinct model assumptions: models H-4.3, H-4.5 and H-4.7 correspond to initial thermal profiles with different central temperatures at the time of the impact, whereas model H-radenv assumes a proto-Jupiter with a radiative envelope. Model H-4.5-lowα uses a shorter mixing length, model H-4.5-semiconv allows for semi-convective mixing, and in model H-4.5-rock the heavy elements are represented by rock instead of water for EOS purposes. The inset zooms in on the region with a normalized radius between 0.15 and 0.5. See text and Extended Data Table 2 for further details.



Extended Data Fig. 8 | Density versus normalized radius for the oblique collision after 4.56 Gyr of evolution. The colours correspond to distinct model assumptions: models O-4.3, O-4.5 and O-4.7 correspond to initial thermal profiles with different central temperatures at the time of the impact. Model O-4.5-lowα uses a shorter mixing length, model O-4.5-semiconv allows for semi-convective mixing, and in model O-4.5-rock the heavy elements are represented by rock instead of water for EOS purposes. The inset zooms in on the region with a normalized radius between 0.15 and 0.4. See text and Extended Data Table 2 for further details.

Extended Data Table 1 | Initial orbital semi-major axes for each embryo of our *N*-body simulation suite

k	*q*	*a* [AU]
5.0	1.11	4.19, 4.66, **5.20**, 5.79, 6.45
6.0	1.14	4.01, 4.56, **5.20**, 5.92, 6.74
8.0	1.19	3.68, 4.37, **5.20**, 6.18, 7.35
10.0	1.24	3.37, 4.19, **5.20**, 6.46, 8.02
6.0	1.14	**5.20**, 5.92, 6.74, 7.67, 8.73
10.0	1.24	**5.20**, 6.46, 8.02, 9.95, 12.36

The location of the embryo that grows into a Jupiter in each case is in boldface. Both *k* and *q* measure the orbital tightness (see equations (1) and (2)).

Extended Data Table 2 | Evolutionary models discussed in this work

Name	log T_c [K]	Collision Type	Heavy-element type	α_{mlt}	Semi-convection
H-4.3	4.3	head-on	water	10^{-1}	no
H-4.5	4.5	head-on	water	10^{-1}	no
H-4.7	4.7	head-on	water	10^{-1}	no
H-radenv	-	head-on	water	10^{-1}	no
H-4.5-lowα	4.5	head-on	water	10^{-3}	no
H-4.5-semiconv	4.5	head-on	water	10^{-1}	yes
H-4.5-rock	4.5	head-on	rock	10^{-1}	no
O-4.3	4.3	oblique	water	10^{-1}	no
O-4.5	4.5	oblique	water	10^{-1}	no
O-4.7	4.7	oblique	water	10^{-1}	no
O-4.5-lowα	4.5	oblique	water	10^{-3}	no
O-4.5-semiconv	4.5	oblique	water	10^{-1}	yes
O-4.5-rock	4.5	oblique	rock	10^{-1}	no

We note that the H-radenv Jupiter-evolution model is unique because it is the result of a Jupiter-formation model[26] that accounts for the accretion shock during the runaway gas accretion.

LETTER

https://doi.org/10.1038/s41586-019-1463-1

Imaging magnetic polarons in the doped Fermi–Hubbard model

Joannis Koepsell[1]*, Jayadev Vijayan[1], Pimonpan Sompet[1], Fabian Grusdt[2,3], Timon A. Hilker[1,5], Eugene Demler[2], Guillaume Salomon[1], Immanuel Bloch[1,4] & Christian Gross[1]

Polarons—electronic charge carriers 'dressed' by a local polarization of the background environment—are among the most fundamental quasiparticles in interacting many-body systems, and emerge even at the level of a single dopant[1]. In the context of the two-dimensional Fermi–Hubbard model, polarons are predicted to form around charged dopants in an antiferromagnetic background in the low-doping regime, close to the Mott insulating state[2–7]; this prediction is supported by macroscopic transport and spectroscopy measurements in materials related to high-temperature superconductivity[8]. Nonetheless, a direct experimental observation of the internal structure of magnetic polarons is lacking. Here we report the microscopic real-space characterization of magnetic polarons in a doped Fermi–Hubbard system, enabled by the single-site spin and density resolution of our ultracold-atom quantum simulator. We reveal the dressing of doublons by a local reduction—and even sign reversal—of magnetic correlations, which originates from the competition between kinetic and magnetic energy in the system. The experimentally observed polaron signatures are found to be consistent with an effective string model at finite temperature[7]. We demonstrate that delocalization of the doublon is a necessary condition for polaron formation, by comparing this setting with a scenario in which a doublon is pinned to a lattice site. Our work could facilitate the study of interactions between polarons, which may lead to collective behaviour, such as stripe formation, as well as the microscopic exploration of the fate of polarons in the pseudogap and 'bad metal' phases.

Polarons usually occur in materials with a strong coupling between mobile charge carriers and collective modes of the background, such as phonons, magnons or spinons[1]. Furthermore, these materials often possess exotic properties, such as spin currents (in organic semiconductors)[9], colossal magnetoresistance (in manganites)[10], pseudogaps (in transition-metal oxides) or high-transition-temperature (high-T_c) superconductivity (in copper oxides)[11]. Even though there are many open questions regarding the microscopic description of these phenomena, some of them can be attributed to polarons, whereas others can emerge from multiple interacting polarons[1,12,13]. The most prominent and conceptually simple electronic model for high-T_c copper oxides is the two-dimensional doped Fermi–Hubbard model, in which an interplay between the kinetic energy of doped charge carriers and a magnetic background supports the formation of magnetic polarons at the single-dopant level. The model consists of spin-1/2 fermions hopping on a two-dimensional lattice with nearest-neighbour (NN) hopping amplitude t and on-site repulsion U between opposite spins. At half-filling, the ground state is a Mott insulating state with antiferromagnetic correlations, owing to an effective superexchange spin coupling of $J = 4t^2/U$. Upon hole or particle doping, dopants can lower their kinetic energy by delocalization. However, each hopping process of the dopant alters the spins of the magnetic background (see Fig. 1), which leads to a growing magnetic cost with increasing delocalization. This problem of a single dopant in an antiferromagnetic environment cannot be solved analytically and requires considerable effort to simulate its properties numerically. However, its understanding marks an important starting point for unravelling the physics of the doped Fermi–Hubbard model.

As a consequence of the competition between magnetic and kinetic energy, theoretical calculations of single dopants in the related t–J model predict the formation of a magnetic polaron[2–7], in which the dopant surrounds itself with a local cloud of reduced antiferromagnetic correlations (see Fig. 1a). Spectroscopic measurements of undoped copper oxides have experimentally probed this single-dopant regime. Even though measurements of dispersion or quasiparticle weights are compatible with the formation of polarons[8], a direct microscopic real-space image of such dressed charge carriers at the single-particle level is still lacking. Furthermore, the evolution from individual polarons into the pseudogap or the 'strange metal' phase at higher doping concentrations is still subject to controversy, leading to diverse theoretical approaches[11,14–16].

Quantum gas microscopy has enabled the direct, model-free real-space characterization of strongly correlated quantum many-body systems. In cold-atom lattice simulators[17], this technique has proved its potential to shed light on the Fermi–Hubbard model, including the detection of long-range spin correlations[18], charge and spin transport[19,20] in two dimensions, as well as incommensurate magnetism[21] in one dimension. Employing the full spin and density resolution of our setup[22], we experimentally confirm the presence and internal structure of magnetic polarons in the low-doping regime of a Fermi–Hubbard system. We observe how double occupations (doublons) are surrounded by a local distortion of antiferromagnetic correlations. Similar qualitative features of the spin correlations around the doublon, as measured in the experiment, are predicted by exact diagonalization of the t–J model, as well as an effective theory that models the polaron as a doublon bound to a spinon by a string of reduced antiferromagnetic correlations[7]. By contrast, by confining a doublon to a single lattice site with an optical tweezer, we observe qualitatively different signatures, underlining the necessity of delocalization for polaron formation.

In our experiment we prepared a balanced mixture of the two lowest hyperfine states of ^{6}Li and adiabatically loaded around 70 atoms into an anisotropic two-dimensional square lattice with spacings $(a_x, a_y) = (1.15, 2.3)\,\mu$m and depths $(8.6E_r^x, 3E_r^y)$, where $E_r^i = h^2/8ma_i^2$ is the recoil energy of the respective lattice, m is the atomic mass and h is the Planck constant. The system is well described by the two-dimensional Fermi–Hubbard model with approximately equal tunnelling amplitudes t_i in both directions, $t_y/h \approx t_x/h = 170$ Hz (see Methods). We tuned the interaction U by using the broad Feshbach resonance in ^{6}Li, such that $U/t_i = 14(1)$, leading to a superexchange coupling of $J/h = 50(5)$ Hz. All uncertainties reported here denote one standard deviation of the mean. We estimate the temperature T of the system to be $k_BT/t = 0.45^{+3}_{-1}$, where k_B is the Boltzmann constant (see Methods). In our study, we used doublon instead of hole doping, because doublons are trapped by our confining potential, avoiding contamination of the

[1]Max-Planck-Institut für Quantenoptik, Garching, Germany. [2]Department of Physics, Harvard University, Cambridge, MA, USA. [3]Department of Physics, Technical University of Munich, Garching, Germany. [4]Fakultät für Physik, Ludwig-Maximilians-Universität, Munich, Germany. [5]Present address: Cavendish Laboratory, University of Cambridge, Cambridge, UK.
*e-mail: joannis.koepsell@mpq.mpg.de

signal by holes created during the detection. A shown in Fig. 1, we realized separate settings in which doublons were allowed to hop between sites (Fig. 1a, left) or were pinned to a single lattice site (Fig. 1a, right). Mobile doublons were prepared using an increased chemical potential, resulting in delocalized doublons in the centre of our harmonically confined lattice with a trapping frequency of about $\omega/(2\pi) = 250$ Hz. For the preparation of immobile doublons we used a tightly focused laser beam (tweezer) at 702 nm with a waist of about 0.5 μm to form a deep attractive local potential. By shining the tweezer with appropriate intensity onto a single lattice site, the deep potential leads to an artificially created trapped doublon at that site (see Fig. 1a). Our detection method[22] allows us to simultaneously reconstruct the local spin and density within a single snapshot (see Fig. 1d). In this way, we can separate the spin and density sectors by measuring local spin correlations between two sites at positions $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ that are singly occupied (indicated by the filled circles below).

$$C(\boldsymbol{r}_1, \boldsymbol{r}_2) = 4\langle S^z_{\boldsymbol{r}_1} S^z_{\boldsymbol{r}_2}\rangle_{\bullet\boldsymbol{r}_1\bullet\boldsymbol{r}_2} \quad (1)$$

We define the value of $C(\boldsymbol{r}_1, \boldsymbol{r}_2)$ as the bond strength between $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$. The spin environment around doublons can be investigated with a three-point doublon–spin correlator, which measures the two-point spin correlation as a function of a detected doublon (indicated by two filled circles) at a third position, $\boldsymbol{r}_0$

$$\begin{aligned} C(\boldsymbol{r}_0;\boldsymbol{r}_1,\boldsymbol{r}_2) &= 4\langle S^z_{\boldsymbol{r}_1} S^z_{\boldsymbol{r}_2}\rangle_{\vdots\boldsymbol{r}_0\bullet\boldsymbol{r}_1\bullet\boldsymbol{r}_2} \equiv C(\boldsymbol{r}_0;\boldsymbol{r},\boldsymbol{d}) \\ &= 4\left\langle S^z_{\boldsymbol{r}_0+\boldsymbol{r}-\frac{\boldsymbol{d}}{2}} S^z_{\boldsymbol{r}_0+\boldsymbol{r}+\frac{\boldsymbol{d}}{2}}\right\rangle_{\vdots\boldsymbol{r}_0\bullet\boldsymbol{r}_0+\boldsymbol{r}-\frac{\boldsymbol{d}}{2}\bullet\boldsymbol{r}_0+\boldsymbol{r}+\frac{\boldsymbol{d}}{2}} \end{aligned} \quad (2)$$

Here, the correlator is expressed in terms of the bond length $\boldsymbol{d} = \boldsymbol{r}_2 - \boldsymbol{r}_1$ of the spin correlation and the bond distance $\boldsymbol{r} = [(\boldsymbol{r}_1 + \boldsymbol{r}_2)/2] - \boldsymbol{r}_0$ from the doublon. This three-point correlator can be understood as defining the origin in each snapshot as the position of a detected doublon and calculating arbitrary two-point spin correlations as a function of distance from that point. For a magnetic polaron, this correlator is expected to reveal the strongly altered spin correlations in the immediate vicinity of doublons (that is, for small bond distances $\boldsymbol{r}$). The remainder of this article will focus on the analysis of $C(\boldsymbol{r}_0; \boldsymbol{r}, \boldsymbol{d})$ for NN ($|\boldsymbol{d}| = 1$), diagonal ($|\boldsymbol{d}| = 1.4$) and next-nearest-neighbour (NNN; $|\boldsymbol{d}| = 2$) spin correlations as a function of bond distance $\boldsymbol{r}$ from the doublons. Even in the Mott insulating regime without doping, quantum fluctuations of doublon–hole pairs lead to a constant background of detected doublons. To distinguish between doped particles and such naturally occurring fluctuations in our signal, we neglect double occupations with holes as NNs (see Methods).

Our access to three-point correlations allows us to study the local distortion of magnetic correlations surrounding doublons and therefore the inner structure of a polaron. Spins located close to a mobile doublon will be affected the strongest by doublon delocalization. Hence, the largest signal is expected for correlations between the four spins that are direct neighbours of a detected doublon; those are diagonal and NNN correlations. The NN correlations closest to the doublon, by contrast, exhibit a larger bond distance and are less sensitive to polaronic spin distortion. Therefore, we first consider the effect of doping on diagonal spin correlations and analyse the correlator defined in equation (2) to evaluate spin correlations as a function of bond distance $\boldsymbol{r}$ from doublons.

To study the doped system, we set the chemical potential such that a doped region of 5 × 3 sites forms with 1.95(1) doublons per experimental realization on average (see Fig. 2a). For each experimental snapshot, doublons are detected at different positions $\boldsymbol{r}_0$. We average the correlator of equation (2) over all positions in the doped region and obtain the average spin correlation around a single doublon $C(\boldsymbol{r}, \boldsymbol{d})$, as displayed in Fig. 2b for diagonal correlations ($|\boldsymbol{d}| = 1.4$). Remarkably, we observe the dressing of doublons with a spin disturbance, which

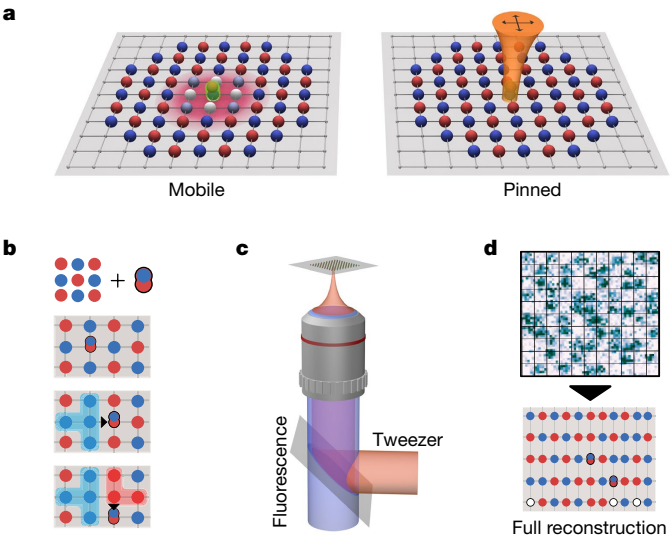


Fig. 1 | Mobile and immobile dopants with ultracold atoms. a, We experimentally study two-dimensional Fermi–Hubbard systems containing fermionic spin-up and spin-down particles (red and blue spheres, respectively), with mobile (left) or immobile (right) doublons (green), using a quantum gas microscope. In the mobile case, antiferromagnetic correlations close to the doublon are diminished (pink shading). This effect is absent in the immobile case, when the doublon is pinned with a focused attractive laser beam (orange). **b**, The hopping of a doublon (black arrow) in the antiferromagnetic background of the Fermi–Hubbard model around half-filling leads to a distorted spin order. With increasing delocalization, antiferromagnetically aligned spin pairs are turned into ferromagnetic ones (red and blue shading, highlighting ferromagnetic regions of different magnetization). As a consequence of this competition, theoretical and experimental evidence points towards polaron formation (see text). **c**, To create immobile localized doublons, we focus an attractive laser beam (orange) onto a single lattice site through the microscope, which collimates the fluorescence light (blue) for detection. **d**, Each captured image corresponds to a projected many-body quantum state. By employing our local Stern–Gerlach technique, we fully resolve spin and density (including holes, represented by white circles), enabling local investigation of the spin environment around doublons. As indicated in the reconstruced image, the Fermi–Hubbard model is implemented with equal tunnelling amplitudes, $t_y = t_x$, but unequal lattice spacings, $a_y = 2a_x$, to allow spin-resolved detection.

confirms the picture of a magnetic polaron. The strong effect on the magnetic correlations is even more pronounced in NNN correlations ($|\boldsymbol{d}| = 2$) across doublons, which reverse their sign with an amplitude a factor of two larger than that of diagonal correlations (see Fig. 2c). Numerical studies at low temperature have found that NNN spin correlations across dopants reverse their sign and become negative[23,24], which has been interpreted as local spin–charge separation and a building block of incommensurate magnetism in two dimensions[24,25]. Our results show that this effect persists even at elevated temperatures. In a frozen-spin picture, this sign reversal can be understood from a single displacement of the doublon (see Fig. 1b), which turns NN spins into NNN ones and thus automatically mixes a strong negative NN signal into the otherwise positive, but weaker, NNN correlations. A similar reasoning also applies to the sign reversal of diagonal spin correlations. Because antiferromagnetic NN correlations are stronger than any other spin correlation even at zero temperature, the string model intuitively predicts this sign flip to be robust also at lower temperatures. For a high enough density of dopants, even local two-point diagonal spin correlations (see equation (1)) can reverse their sign, as shown in Methods and reported in refs [26–28].

In addition, we analysed the correlations between doublons. At our temperature they appear anti-correlated at short distances and uncorrelated otherwise within our measurement uncertainty (see Methods). This is in agreement with other recent observations[28]

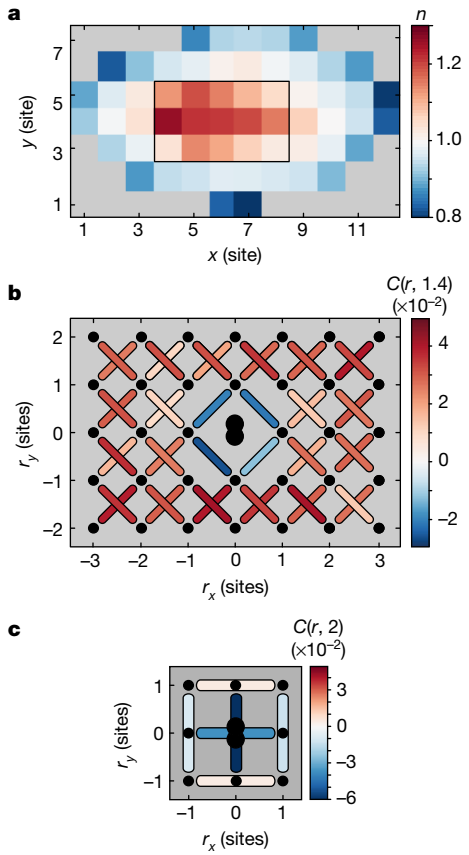


Fig. 2 | Mobile doublons dressed by local spin disturbance. a, Density (n) distribution for mobile doublons. In the doped region (inner black box), two doublons on average delocalize in an area of 5 × 3 sites. **b**, Diagonal spin correlations, represented by bonds connecting two sites (black dots) and sorted according to their distance from doublons (double black circle at centre). Correlations are negative only in the immediate vicinity of a doublon and positive farther away. **c**, NNN spin correlations across and next to detected doublons. As in the case of diagonal correlations (**b**),the correlations across doublons are sign-flipped with respect to the antiferromagnetic background value. Our experimental results confirm the formation of a magnetic polaron, in which doublons are dressed by a local spin distortion (see Fig. 1a, left).

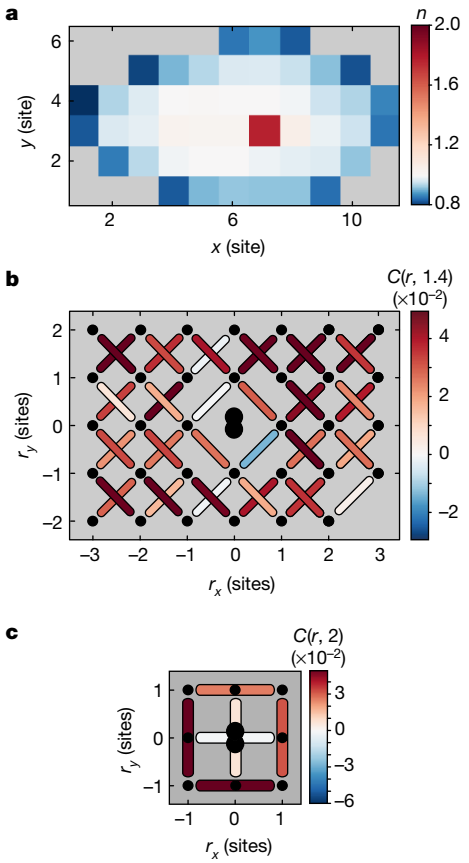


Fig. 3 | Spin correlations around trapped doublons. a, Density distribution for pinned doublons. An attractive laser beam (tweezer; 702 nm) focused on a single site artificially increases the density in an undoped system at a specific site to about 1.77(1). **b**, Diagonal spin correlations around doublons trapped in the tweezer. The sign-flipped spin distortion vanishes, in contrast to the mobile case. **c**, NNN spin correlations across and next to pinned doublons. Although spins across the trapped doublon are uncorrelated, correlations neighbouring the trapped doublon are slightly enhanced compared to the background value (see Fig. 4c). Trapping doublons with a tweezer beam prevents the competition between kinetic and magnetic energy and suppresses polaron formation (see Fig. 1a, right).

and supports our treatment of each doublon as an independent fermionic particle.

To demonstrate that the mobility of the doublon is key for polaron formation, we now investigate the effect of an artificially introduced localized doublon on the surrounding magnetic correlations. We set the chemical potential so as to prepare a system without doping and we adiabatically ramp up the power of an optical tweezer focused on a single central site while simultaneously ramping up the lattice. The final tweezer depth is set such that the density of that site saturates at 1.77(1) (see Fig. 3a). We do not achieve a perfectly deterministic doublon preparation in our experiments, probably because of detection errors and higher-band effects (see Methods). We analyse the same doublon-conditioned three-point correlator $C(\boldsymbol{r}_0; \boldsymbol{r}, \boldsymbol{d})$ for diagonal spin correlations, as before, with $\boldsymbol{r}_0$ fixed to the pinned site (see Fig. 3b). As expected, the strong spin distortion and, most importantly, the sign reversal of correlations is absent in this case. Instead, magnetic correlations across the trapped site are only moderately reduced compared to the undoped background (see Fig. 3c).

To enable a quantitative study and a comparison to theoretical models, we group the three-point spin correlations by the magnitude of their bond distance $r = |\boldsymbol{r}|$ from doublons. Measured NN, diagonal and NNN spin correlations are shown in Fig. 4. The local distortion of spin correlations around mobile doublons is visible in all correlators. Sign reversal of diagonal and NNN correlations occurs at a mean bond distance of one site, yielding a diameter (and estimated polaron size) of around two lattice sites. We compare our findings to theoretical model calculations carried out for the estimated temperature of our system (see Fig. 4). For the mobile case, an effective string model of magnetic polarons is used, assuming frozen spin dynamics[7]. Remarkably, similar amplitude changes of correlations, and hence a similar polaron radius, is predicted (also found in the exact diagonalization results of the t–J model on a 4 × 4 system; see Supplementary Information). Furthermore, the sign changes of correlations in the vicinity of the doublon are reproduced in this model, as seen also in Fig. 4e, f. Quantitative differences between the effective model and the experiment remain; however, this is expected owing to the moderate separation of spin and hole dynamics ($J/t = 0.3$) and the elevated temperatures in the experimental system. In the case of pinned doublons, two effects can be observed. First, the sign flip of correlations observed in the mobile case vanishes and the closest-distance diagonal and NNN spins appear uncorrelated; this is captured by an exact diagonalization calculation of the t–J model with zero tunnelling of the excess doublon (see Fig. 4). This can be explained by the fact that the doublon effectively blocks a path linking the spins next to it and prevents them from building up a correlation, given the finite temperature. Second, an enhancement of certain spin correlations around the pinned site is visible in the closest NN correlations and distance-1 NNN correlations. This effect is expected from

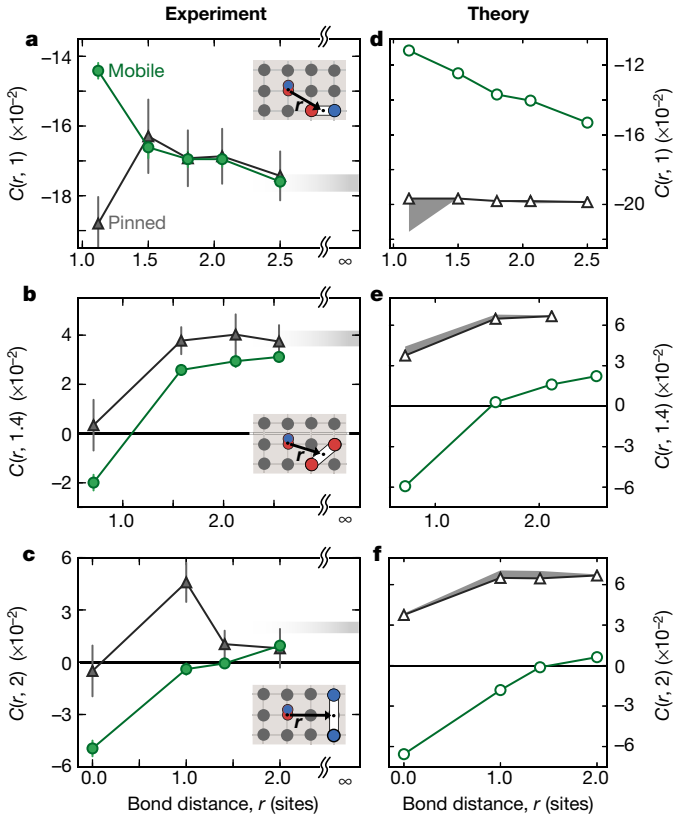


Fig. 4 | Spin correlations as a function of bond distance from doublons. **a**–**f**, Comparison between experiment (**a**–**c**) and numerical calculations performed using a string model for magnetic polarons and exact diagonalization of an immobile doublon in the *t*–*J* model (**d**–**f**). NN (**a**, **d**), diagonal (**b**, **e**) and NNN (**c**, **f**) spin correlations as a function of bond distance from mobile (green) or immobile (black) doublons. The insets show one examplary bond (white) between two particles and its distance ***r*** from a doublon (double circle). Error bars denote one standard error of the mean (s.e.m.). For mobile doublons, diagonal and NNN correlations within an average bond distance of one lattice site are sign-flipped with respect to the antiferromagnetic background. Correlations quickly recover at larger distances to a value approaching the undoped antiferromagnetic value, represented by the grey band at distance ∞ with a width of 2 s.e.m. The string model (green in **d**–**f**) predicts similar correlation changes with bond distance, as well as the sign reversal of diagonal and NNN correlations. For pinned doublons, the diagonal and NNN correlations at the smallest bond distance are not sign-flipped and the polaronic distortion is strongly reduced. The amplitude of the remaining weakening of the magnetic correlations is consistent with exact diagonalization calculations of a trapped doublon (black in **d**–**f**). Owing to the finite tweezer size (see text), a slight enhancement of NN and NNN correlations around a distance of one site is visible in the experiment, which can be captured by exact diagonalization with 10% enhanced spin exchange on neighbouring sites (grey band).

local energy shifts around the pinned site that are caused by the finite extent of the optical-tweezer beam (see Methods). Those energy shifts lead to a locally enhanced superexchange coupling *J* and can therefore cause stronger correlations. Exact diagonalization calculations at finite temperature with up to 10% increased superexchange coupling in the vicinity reproduce our experimental correlation enhancement (see Methods). Nonetheless, correlations across the doublon and the shortest-distance diagonal correlations are almost unaffected by this small systematic enhancement.

We have presented single-particle-resolution imaging of a magnetic polaron in a doped Fermi–Hubbard system by revealing the dressing of mobile doublons with a spin distortion. We identified a compact polaron size of about two sites and characterized its inner structure, in which spin correlations can exhibit even sign reversal compared to the undoped system. Our findings qualitatively agree with numerical predictions. Artificially localizing the doublon considerably reduces the spin distortion and the sign flip disappears, as the competition between kinetic and magnetic energy is suppressed. The ability to spatially resolve the dressing cloud of polarons enables a fundamentally new approach to experimentally characterizing such quasiparticles at the microscopic level and could be applied to study the polaron physics of impurities immersed in bosonic[29] or fermionic gases[30] and provide observables for the exploration of strongly correlated phenomena and their microscopic origin. In the future, the effective mass or the quasiparticle weight of the polaron could be probed by transport[19,20] or spectroscopic methods[31]. By implementing larger and more homogeneous systems, as well as new cooling schemes[18,32,33], a microscopic study of polaron–polaron interactions and the crossover from polarons to the emergence of pseudogap, strange-metal and stripe phases or pairing is within reach.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1463-1.

Received: 16 November 2018; Accepted: 10 June 2019;
Published online 14 August 2019.

1. Alexandrov, S. & Devreese, J. T. *Advances in Polaron Physics* (Springer, 2010).
2. Schmitt-Rink, S., Varma, C. M. & Ruckenstein, A. E. Spectral function of holes in a quantum antiferromagnet. *Phys. Rev. Lett.* **60**, 2793–2796 (1988).
3. Shraiman, B. I. & Siggia, E. D. Mobile vacancies in a quantum Heisenberg antiferromagnet. *Phys. Rev. Lett.* **61**, 467–470 (1988).
4. Sachdev, S. Hole motion in a quantum Néel state. *Phys. Rev. B* **39**, 12232–12247 (1989).
5. Kane, C. L., Lee, P. A. & Read, N. Motion of a single hole in a quantum antiferromagnet. *Phys. Rev. B* **39**, 6880–6897 (1989).
6. Dagotto, E., Moreo, A. & Barnes, T. Hubbard model with one hole: ground-state properties. *Phys. Rev. B* **40**, 6721–6725 (1989).
7. Grusdt, F. et al. Parton theory of magnetic polarons: mesonic resonances and signatures in dynamics. *Phys. Rev. X* **8**, 011046 (2018).
8. Schrieffer, J. R. *Handbook of High-Temperature Superconductivity* (Springer, 2007).
9. Watanabe, S. et al. Polaron spin current transport in organic semiconductors. *Nat. Phys.* **10**, 308–313 (2014).
10. Ramirez, A. P. Colossal magnetoresistance. *J. Phys. Condens. Matter* **9**, 8171–8199 (1997).
11. Lee, P. A., Nagaosa, N. & Wen, X.-G. Doping a Mott insulator: physics of high-temperature superconductivity. *Rev. Mod. Phys.* **78**, 17–85 (2006).
12. Trugman, S. A. Interaction of holes in a Hubbard antiferromagnet and high-temperature superconductivity. *Phys. Rev. B* **37**, 1597–1603 (1988).
13. Schrieffer, J. R., Wen, X. & Zhang, S. C. Dynamic spin fluctuations and the bag mechanism of high-T_c superconductivity. *Phys. Rev. B* **39**, 11663–11679 (1989).
14. Anderson, P. W. The resonating valence bond state in La_2CuO_4 and superconductivity. *Science* **235**, 1196–1198 (1987).
15. Auerbach, A. & Larson, B. E. Small-polaron theory of doped antiferromagnets. *Phys. Rev. Lett.* **66**, 2262–2265 (1991).
16. Punk, M., Allais, A. & Sachdev, S. Quantum dimer model for the pseudogap metal. *Proc. Natl Acad. Sci. USA* **112**, 9552–9557 (2015).
17. Gross, C. & Bloch, I. Quantum simulations with ultracold atoms in optical lattices. *Science* **357**, 995–1001 (2017).
18. Mazurenko, A. et al. A cold-atom Fermi–Hubbard antiferromagnet. *Nature* **545**, 462–466 (2017).
19. Brown, P. T. et al. Bad metallic transport in a cold atom Fermi–Hubbard system. *Science* **363**, 379–382 (2019).
20. Nichols, M. A. et al. Spin transport in a Mott insulator of ultracold fermions. *Science* **363**, 383–387 (2019).
21. Salomon, G. et al. Direct observation of incommensurate magnetism in Hubbard chains. *Nature* **565**, 56–60 (2019); correction 566, E5 (2019).
22. Boll, M. et al. Spin- and density-resolved microscopy of antiferromagnetic correlations in Fermi–Hubbard chains. *Science* **353**, 1257–1260 (2016).
23. White, S. R. & Scalapino, D. Hole and pair structures in the t–J model. *Phys. Rev. B* **55**, 6504–6517 (1997).
24. Martins, G. B., Eder, R. & Dagotto, E. Indications of spin-charge separation in the two-dimensional t–J model. *Phys. Rev. B* **73**, 170–173 (1999).
25. Martins, G. B., Gazza, C., Xavier, J. C., Feiguin, A. & Dagotto, E. Doped stripes in models for the cuprates emerging from the one-hole properties of the insulator. *Phys. Rev. Lett.* **84**, 5844–5847 (2000).
26. Parsons, M. F. et al. Site-resolved measurement of the spin-correlation function in the Fermi–Hubbard model. *Science* **353**, 1253–1256 (2016).

27. Cheuk, L. W. et al. Observation of spatial charge and spin correlations in the 2D Fermi–Hubbard model. *Science* **353**, 1260–1264 (2016).
28. Chiu, C. S. et al. String patterns in the doped Hubbard model. Preprint at https://arxiv.org/abs/1810.03584 (2018).
29. Devreese, J. T. & Alexandrov, A. S. Froehlich polaron and bipolaron: recent developments. *Rep. Prog. Phys.* **72**, 066501 (2009).
30. Radzihovsky, L. & Sheehy, D. E. Imbalanced Feshbach-resonant Fermi gases. *Rep. Prog. Phys.* **73**, 076501 (2010).
31. Stewart, J. T., Gaebler, J. P. & Jin, D. S. Using photoemission spectroscopy to probe a strongly interacting Fermi gas. *Nature* **454**, 744–747 (2008).
32. Lubasch, M., Murg, V., Schneider, U., Cirac, J. I. & Bañuls, M.-C. Adiabatic preparation of a Heisenberg antiferromagnet using an optical superlattice. *Phys. Rev. Lett.* **107**, 165301 (2011).
33. Kantian, A., Langer, S. & Daley, A. J. Dynamical disentangling and cooling of atoms in bilayer optical lattices. *Phys. Rev. Lett.* **120**, 060401 (2018).

METHODS

Experimental sequence. The preparation of cold Fermi–Hubbard systems closely followed the procedure described in Salomon et al.[21]. We started by preparing a balanced mixture of the two lowest hyperfine states of fermionic ^{6}Li (hyperfine and magnetic quantum numbers of $F = 1/2$ and $m_F = \pm 1/2$), which was harmonically confined in a single two-dimensional plane of a $40E_r^z$-deep optical lattice with 3.1 μm spacing in the z direction. The final atom number was set by the evaporation parameters. Subsequently, we ramped the depth of the x and y lattices with spacings of $a_x = 1.15$ μm and $a_y = 2.3$ μm linearly from $0E_r^i$ to their final values of $8.6E_r^x$ and $3E_r^y$ within 210 ms. The final lattice depths were optimized with undoped Mott insulators to give strong and isotropic spin correlations. From band-width calculations we extracted the tight-binding NN tunnelling amplitudes $t_x/h = 170$ Hz and $t_y/h = 180$ Hz. The NNN tunnelling amplitude along the y direction was below $0.1t_y$. Using the broad Feshbach resonance of ^{6}Li, the scattering length was tuned during the lattice ramp from $350a_B$ to $2150a_B$ in two linear ramps, where a_B denotes the Bohr radius. The first one ramped to $980a_B$ within 150 ms and the second ramp increased the scattering length to $2150a_B$ in 60 ms. By applying a local Stern–Gerlach detection technique[22] and subsequent Raman sideband cooling in a pinning lattice[34], the local spin and density of each lattice site was obtained with an average fidelity of 97%.

Data analysis. The work presented here used a dataset for pinned doublons and one for mobile doublons consisting of 33,669 and 9,002 images, respectively. In the analysis, we considered only shots with total spin $|S^z_{tot}| \leq 3.5$, in order to filter out fluctuations in our spin detection scheme[21] and strongly magnetized clouds. This corresponds to a maximum allowed magnetization of $|S^z_{tot}|/N \approx 0.05$ and approximately 68% of the images recorded. All further analysis was performed on lattice sites with a mean density of $n \geq 0.7$. This region corresponds to the sites shown in Figs. 2a, 3a for the mobile and pinned cases, respectively. To exclude clouds heated from inelastic three-body collisions, images with a total number of holes of $\sum_{ROI} n_h \geq 8$ contained in this region of interest were discarded (which amounts to neglecting around 16% of the data). Computing the doublon–hole correlation function $g_2 = [\langle \hat{n}_d \hat{n}_h \rangle / (\langle \hat{n}_d \rangle \langle \hat{n}_h \rangle)] - 1$ reveals doublon–hole bunching at NN distances (see Extended Data Fig. 1), which indicates the presence of doublon–hole fluctuations. To distinguish between doped excess doublons and doublon–hole fluctuations, we excluded doublons with holes as NNs from the analysis. The resulting dataset statistics after processing is shown in Extended Data Fig. 2 for the mobile case.

Doping calibration. To control the doping, we measured the number of double occupations per number of atoms (N_{doub}/N) in our system as a function of the mean atom number N (see Extended Data Fig. 3a), which is set by our final evaporation parameters. For low atom numbers, the doublon fraction saturates below 4%, which we attribute to quantum fluctuations in the form of doublon–hole pairs. The background of doublon–hole fluctuations is confirmed by discarding doublons with holes as NNs, obtaining the curve of doped doublons versus total atom number shown in Extended Data Fig. 3b. For low atom numbers, no doped doublons are present, whereas at higher atom numbers finite doping sets in. To probe individual mobile doublons in the Fermi–Hubbard model, we used systems with about 72 atoms and around two doped doublons. To study the effect of localized doublons created using an optical tweezer (see below), we used smaller systems with around 55 atoms to avoid the effect of doping.

Tweezer depth calibration. For the pinned doublon case, we ramped the power of an optical tweezer beam focused on a single lattice site to its final value simultaneously with the x and y lattice depth. The tweezer depth was calibrated in a separate measurement by determining the density of the target site as a function of the final tweezer power. As shown in Extended Data Fig. 4, the density first increases with power and then saturates below 1.8, independent of higher final powers. The total detected local density n of 1.77 is $n = 3 \times n_t + 2 \times n_d + 1 \times n_s + 0 \times n_h$, where n_t, n_d, n_s and n_h are the triplon, doublon, singlon and hole density, respectively. For our measurement of localized doublons, we set the tweezer depth to the value at which the density starts saturating. At this tweezer power, the hole density at that site is $n_h = 0.07$, the singlon density is $n_s = 0.13$, the doublon density is $n_d = 0.74$ and a small triplon density of $n_t = 0.05$ exists, which we attribute to imperfections in the detection of doublons, imperfections in the loading procedure and coupling of the y direction to higher bands. In combination with our finite imaging fidelity of 97%, this explains why a deterministic preparation of the doublon is not fully achieved.

Tweezer effect on neighbouring sites. Considering our experimental point-spread function and our numerical aperture of 0.5, the intensity of the 702-nm light radially falls off to 30% at a distance of 600 nm from the maximum. Our point-spread function furthermore shows two asymmetric distorted side-maxima with around 10% intensity, and imperfections in our compensation of the chromatic focal shift between our imaging light at 671 nm and the tweezer beam at 702 nm will lead to a finite-energy shift on neighbouring sites. The total tweezer depth can be approximated by the interaction energy U, and we expect that neighbouring sites (especially in the x direction) are shifted in energy by up to $0.3U$. Such a detuning alters the spin exchange between those sites according to $J = 2t^2[1/(U + \Delta) + 1/(U - \Delta)]$ (ref. [35]). As mentioned in the main text, this explains the enhancement of certain spin correlations around the pinned site and is consistent with the slightly increased average densities on the left (right) of the pinned site in the x direction to 1.024(5) (1.059(5)). When modelling this effect in exact diagonalization calculations, an enhanced spin exchange of 10% between all eight sites surrounding the pinned doublon was assumed.

Temperature estimation. To estimate the temperature of the clouds, we compared the loss-corrected NN spin correlations $\langle S^z_r S^z_{r+e_i} \rangle$, where $e_i = \{e_x, e_y\}$, close to half-filling with numerical linked-cluster expansions up to ninth order for homogeneous systems[36]. We used Wynn's algorithm[37] to sum the terms of the series and obtain NN spin correlations as a function of the density for $U/t = 13$. This value is the lowest estimate of the interaction strength U and takes into account the renormalization for low lattice depths[38]. The experimental spin correlations as a function of density were obtained by averaging over sites with local densities between 0.9 and 1.1 in bins ranging from 0.02 to 0.04 to collect enough statistics. We find that our experimental correlations compare well with numerical linked-cluster expansion results at a temperature of $T \in [0.43t, 0.46t]$ (see Extended Data Fig. 5). To account for the uncertainty in the exact interaction U, we conservatively estimate our temperature to be $k_B T/t = 0.45^{+3}_{-1}$ (see Extended Data Fig. 5). Neither the experimental nor the numerical results have a substantial dependence on temperature with respect to such temperature changes.

Doublon–doublon correlations. Possible interactions between doped excess doublons can be detected by the normalized density–density g_2 correlation function $g_2(\boldsymbol{r}_1, \boldsymbol{r}_2) = [\langle \hat{n}_d(\boldsymbol{r}_1) \hat{n}_d(\boldsymbol{r}_2) \rangle / (\langle \hat{n}_d(\boldsymbol{r}_1) \rangle \langle \hat{n}_d(\boldsymbol{r}_2) \rangle)] - 1$. Here, the operator $\hat{n}_d$ measures the density of real excess doublons without doublon–hole fluctuations. As seen in Extended Data Fig. 6, doublons are anti-correlated at short distances and quickly become uncorrelated within our measurement precision. The anti-correlation is expected for free fermions and our current statistical uncertainty does not allow us to resolve possible small interaction effects at the realized temperature.

Extended polaron analysis. The spin correlator $C(\boldsymbol{r}, \boldsymbol{d})$ discussed in the main text was used to determine the polaronic spin environment of mobile doublons. This correlator is the result of averaging $C(\boldsymbol{r}_0; \boldsymbol{r}, \boldsymbol{d})$ over the positions $\boldsymbol{r}_0$ of mobile doublons. Here we show that the spin distortion is consistently dressing the doublon, independently of the position in the trap. We study the NN, diagonal and NNN correlations with the shortest bond distance $\boldsymbol{r}$ to the doublon as a function of position $\boldsymbol{r}_0$. To maintain a sufficiently high signal-to-noise ratio, we average bonds isotropically. Furthermore, we contrast this to the case in which the position $\boldsymbol{r}_0$ is singly occupied instead

$$C_{singlon}(\boldsymbol{r}_0; \boldsymbol{r}, \mathbf{d}) \equiv C_{singlon}(\boldsymbol{r}_0; \boldsymbol{r}_1, \boldsymbol{r}_2) = 4 \langle S^z_{r_1} S^z_{r_2} \rangle_{\bullet \boldsymbol{r}_0 \bullet \boldsymbol{r}_1 \bullet \boldsymbol{r}_2} \quad (3)$$

The total correlation strength for those two different cases is shown in Extended Data Fig. 7 as a function of position $\boldsymbol{r}_0$ in the system for NN, diagonal and NNN correlations. A strong difference in the local spin environment is observed, depending on whether a doublon or singlon is present at a specific site. The spin distortion that dresses the doublon is strongest for NNN correlations and weakest for NN correlations, which can be understood by considering that NNN correlations are much closer (bond distance 0) to the doublon than NN correlations (bond distance of 1.1). When a singlon occupies a certain position, the strong spin distortion is absent. In this case, the spin correlation surrounding the singlon does not fully return to the background value of an undoped system, because polarons are still present in the system and their average distance from the singlon is of the order of one to two lattice sites. This is also responsible for the varying correlation strength of singlons at different positions. When the singlon is considered in regions of higher density, the average distance to polarons decreases, leading to a parasitic reduction in correlation strength.

Diagonal two-point spin correlations. Two-point spin correlations along the lattice diagonal are shown in Extended Data Fig. 8. In regions with high doublon density (see lattice site positions in Fig. 2a), these two-point correlations flip their sign.

NN spin correlations. Two- and three-point NN spin correlations (equations (1), (2) with $|\boldsymbol{d}| = 1$) are shown in Extended Data Fig. 9. The spin distortion dressing mobile doublons is also visible here. Nonetheless, as explained above, the signal-to-noise ratio is weaker than for the other correlators. In Extended Data Fig. 9c, NN correlations are shown for the case of pinned doublons. The local enhancement of correlations is visible at the closest bond distance.

Data availability

The datasets generated and analysed during this study are available from the corresponding author upon reasonable request.

34. Omran, A. et al. Microscopic observation of Pauli blocking in degenerate fermionic lattice gases. *Phys. Rev. Lett.* **115**, 263001 (2015).

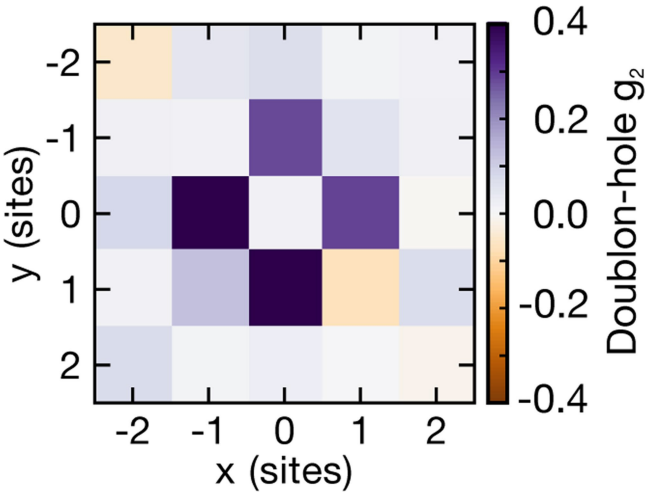
35. Trotzky, S. et al. Time-resolved observation and control of superexchange interactions with ultracold atoms in optical lattices. *Science* **319**, 295–299 (2008).
36. Khatami, E. & Rigol, M. Thermodynamics of strongly interacting fermions in two-dimensional optical lattices. *Phys. Rev. A* **84**, 053611 (2011).
37. Rigol, M., Bryant, T. & Singh, R. R. P. Numerical linked-cluster algorithms. I. Spin systems on square, triangular, and kagomé lattices. *Phys. Rev. E* **75**, 061118 (2007).
38. Büchler, H. P. Microscopic derivation of Hubbard parameters for cold atomic gases. *Phys. Rev. Lett.* **104**, 090402 (2010).

Acknowledgements We thank D. Stamper-Kurn, Y. Wang, E. Manousakis, S. Sachdev and T. Giamarchi for discussions. This work benefited from financial support by the Max Planck Society (MPG), the European Union (UQUAM, FET-Flag 817482, PASQUANS) and the Max Planck Harvard Research Center for Quantum Optics (MPHQ). J.K. gratefully acknowledges funding from Hector Fellow Academy. F.G. and E.D. acknowledge support from Harvard-MIT CUA, NSF grant number DMR-1308435 and AFOSR-MURI Quantum Phases of Matter (grant FA9550-14-1-0035). F.G. acknowledges financial support from the Gordon and Betty Moore Foundation through the EPiQS programme and from the Technical University of Munich – Institute for Advanced Study, funded by the German Excellence Initiative and the European Union FP7 under grant agreement 291763, from DFG grant number KN 1254/1-1 and DFG TRR80 (Project F8).

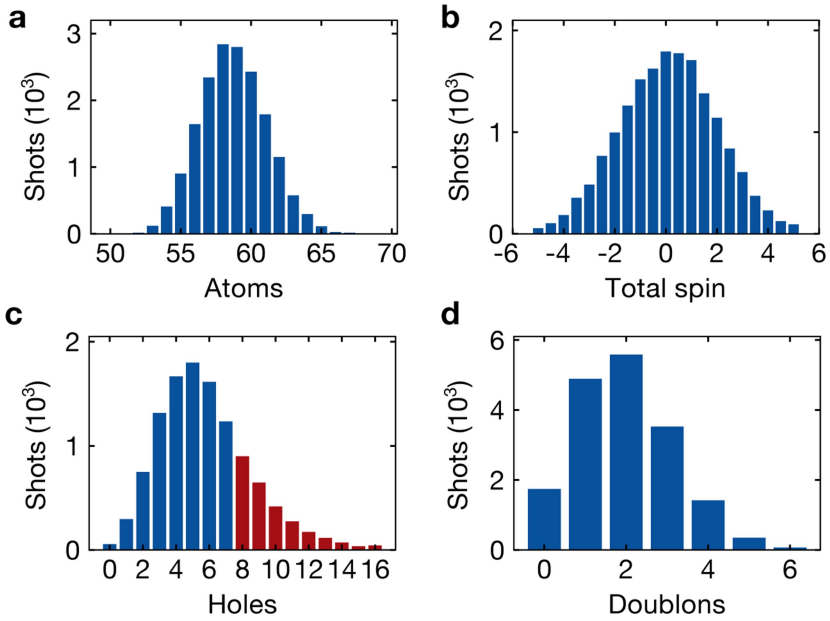
Author contributions J.K., J.V., P.S., T.A.H., G.S., I.B. and C.G. planned the experiment and analysed and discussed the data. F.G. and E.D. performed the theoretical simulations. All authors contributed to the interpretation of the data and writing of the manuscript.

Competing interests The authors declare no competing interests.

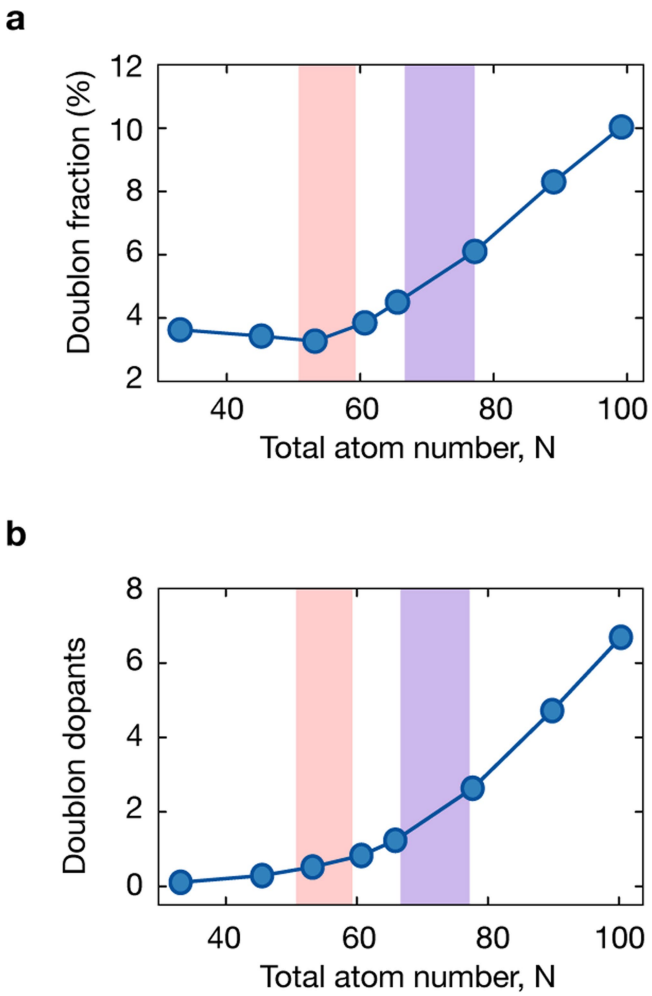
Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1463-1.
Correspondence and requests for materials should be addressed to J.K.



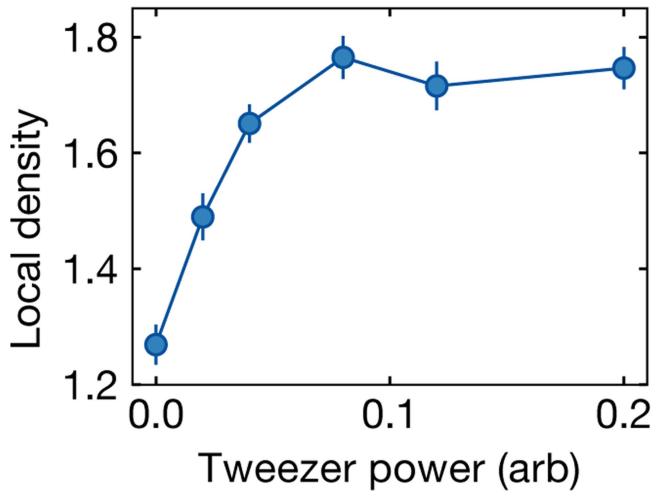
Extended Data Fig. 1 | Doublon–hole correlation. Two-point correlation function g_2 between double occupations and holes, showing a strong bunching effect at NN distances. This motivates us to neglect double occupations with holes as NNs in the analysis of mobile doublons.



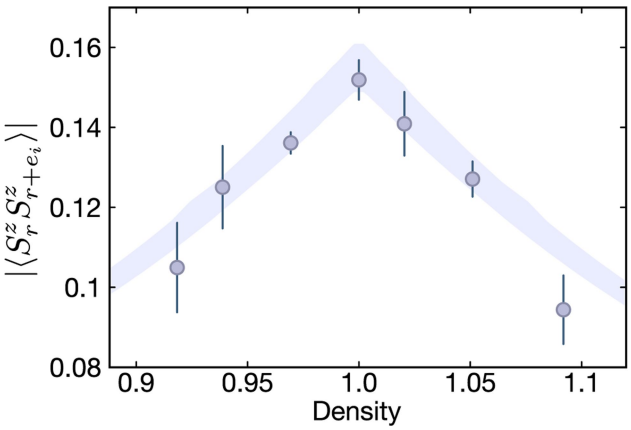
Extended Data Fig. 2 | Dataset statistics for the measurement of mobile doublons. a–c, Distribution of the number of atoms (**a**), spins (**b**) and holes (**c**) in the region with density greater than 0.7. Red bars in **c** indicate shots discarded by the applied hole filter (see text). **d**, Number of mobile doublons (doublon–hole fluctuations subtracted) in the doped 5 × 3-site region.



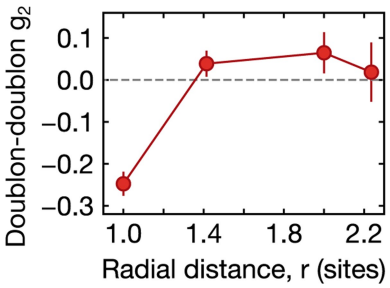
Extended Data Fig. 3 | Doping calibration. a, **b**, When scanning our final evaporation parameters, we measured the fraction of double occupations (**a**)and the number of doped doublons (**b**; excluding doublon–hole fluctuations) in the system as a function of the mean total atom number, *N*. Statistical error bars are smaller than the marker size. Pinned doublon measurements were taken in an undoped system (pink bar). For the mobile doublon dataset, settings for weak doping were used (purple bar). The bar width represents the standard deviation, obtained from atom number fluctuations.



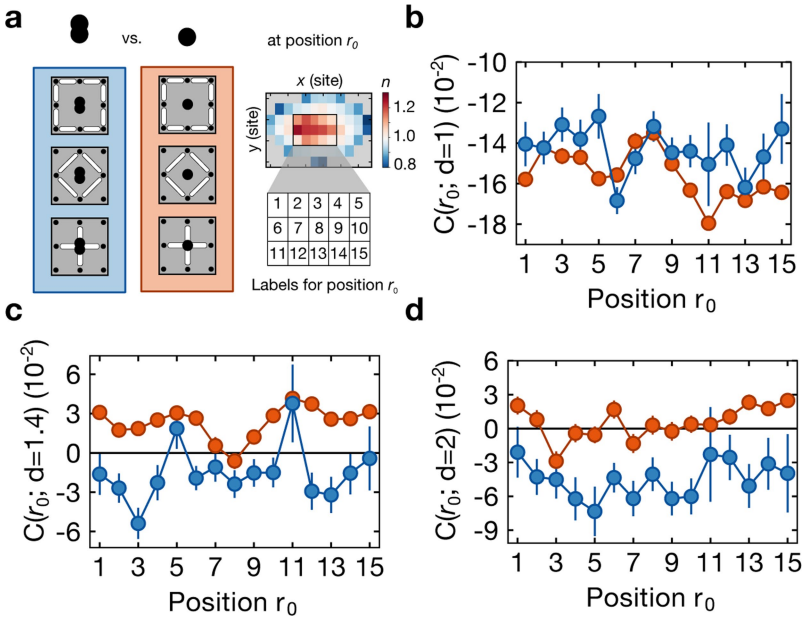
Extended Data Fig. 4 | Calibration of tweezer power. Density of the lattice site on which the tweezer is focused as a function of final tweezer power. Error bars denote one s.e.m. For the realization of pinned doublons the power was set to 0.11 (arbitrary units).



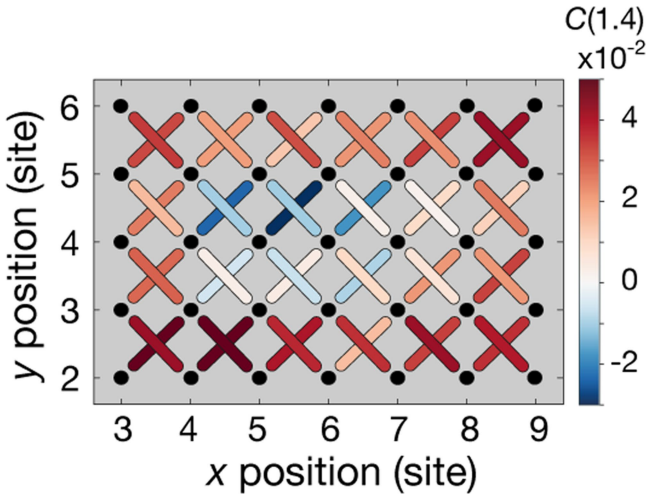
Extended Data Fig. 5 | Temperature estimation. Two-point NN spin correlations as a function of binned density. Error bars denote one s.e.m. Upper (lower) values of the purple band correspond to temperatures of $T/t = 0.43$ (0.46) in numerical linked-cluster expansions at $U/t = 13$.



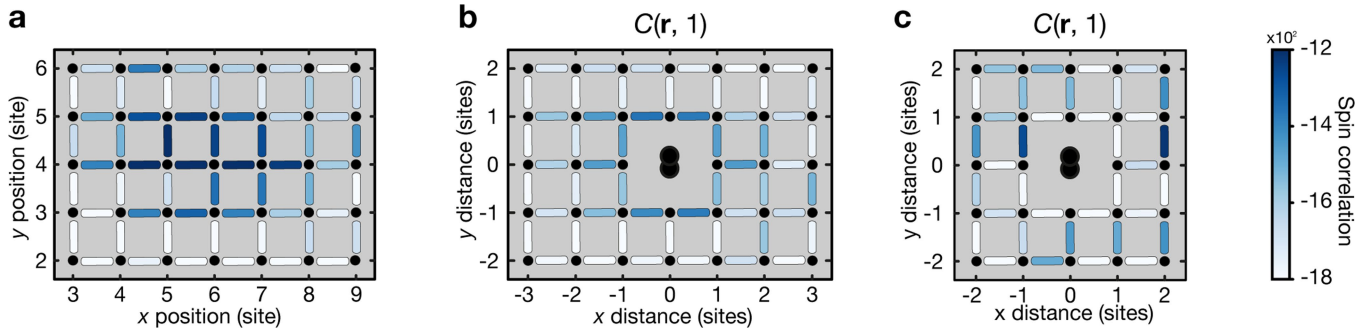
Extended Data Fig. 6 | Radially averaged doublon–doublon correlation function, g_2. In our system, excess doublons appear anti-correlated at short distances and quickly become uncorrelated at longer distances. Error bars denote one s.e.m.



Extended Data Fig. 7 | Extended polaron analysis. a, Comparison of the local spin environment around a lattice site $\boldsymbol{r}_0$ occupied by a doublon (double black circle) or a singlon (single black circle). To simplify the notation, site positions in the doped region of our system are labelled from 0 to 15 (see inset at right). At any position in the 5 × 3-site system, spin distortion is present in NN (**b**), diagonal (**c**) and NNN (**d**) correlations whenever a doublon is detected (blue), and absent whenever a singlon is detected (red). Error bars denote one s.e.m. NNN correlations are measured across doublons and have a high signal-to-noise ratio, which we attribute to their short bond distance of 0, compared to, for example, NN correlations (bond distance of 1.1).



Extended Data Fig. 8 | Diagonal two-point spin correlations. Spin correlations in the central region (see lattice site positions in Fig. 2), represented by bonds connecting the two sites (black dots). At the centre, a clear reduction of correlations from the positive antiferromagnetic background value is visible. In the area of highest doublon density, correlations even flip sign and become negative.



Extended Data Fig. 9 | NN correlations around mobile or pinned doublons. a, Two-point NN correlations in the central region for the mobile doublon setting, represented by colored bonds between lattice sites (black dots). **b**, NN spin correlations as a function of distance from mobile doublons. **c**, NN correlations around pinned doublons. The enhancement effect of correlations is visible in the strong bonds surrounding the trapping site.

LETTER

https://doi.org/10.1038/s41586-019-1428-4

Global entangling gates on arbitrary ion qubits

Yao Lu[1,3]*, Shuaining Zhang[1,3], Kuan Zhang[1,2,3], Wentao Chen[1], Yangchao Shen[1], Jialiang Zhang[1], Jing-Ning Zhang[1] & Kihwan Kim[1]*

Quantum computers can efficiently solve classically intractable problems, such as the factorization of a large number[1] and the simulation of quantum many-body systems[2,3]. Universal quantum computation can be simplified by decomposing circuits into single- and two-qubit entangling gates[4], but such decomposition is not necessarily efficient. It has been suggested that polynomial or exponential speedups can be obtained with global *N*-qubit (*N* greater than two) entangling gates[5–9]. Such global gates involve all-to-all connectivity, which emerges among trapped-ion qubits when using laser-driven collective motional modes[10–14], and have been implemented for a single motional mode[15,16]. However, the single-mode approach is difficult to scale up because isolating single modes becomes challenging as the number of ions increases in a single crystal, and multi-mode schemes are scalable[17,18] but limited to pairwise gates[19–23]. Here we propose and implement a scalable scheme for realizing global entangling gates on multiple $^{171}Yb^{+}$ ion qubits by coupling to multiple motional modes through modulated laser fields. Because such global gates require decoupling multiple modes and balancing all pairwise coupling strengths during the gate, we develop a system with fully independent control capability on each ion[14]. To demonstrate the usefulness and flexibility of these global gates, we generate a Greenberger–Horne–Zeilinger state with up to four qubits using a single global operation. Our approach realizes global entangling gates as scalable building blocks for universal quantum computation, motivating future research in scalable global methods for quantum information processing.

A representative entangling gate with more than two qubits is the global entangling gate, which can generate entanglement among all involved qubits in a symmetric way. A global entangling gate acting on N qubits is defined as

$$\mathrm{GE}_N(\Theta) = \exp\left(-i\Theta\sum_{j<j'}^{N}\sigma_x^j\sigma_x^{j'}\right) \quad (1)$$

where all of the two-body couplings are driven simultaneously with strength Θ, and σ_x^j is the Pauli operator on the jth qubit. A global entangling gate applied to N qubits is equivalent to $N(N-1)/2$ pairwise entangling gates[9], which provides the possibility of simplifying quantum circuits. For example, the N–1 pairwise entangling operations involved in the preparation of the N-qubit Greenberger–Horne–Zeilinger (GHZ) state[20,24] can be replaced by a single global entangling gate $\mathrm{GE}_N(\pi/4)$, as shown in Fig. 1a. In fact, several theoretical works have already indicated that numerous quantum algorithms and universal quantum simulations of various many-body systems would benefit from global entangling gates for the efficient construction of quantum circuits. In particular, a set of $O(N)$ controlled NOT gates in the quantum phase estimation algorithm[9]—as well as each $O(N)$–body interaction term that emerges in the simulation of fermionic systems owing to the Jordan–Wigner transformation[5,6], which requires $O(N)$ pairwise gates—can be efficiently implemented by $O(1)$ global gates. Moreover, because the global gate contains all of the pairwise couplings, we can flexibly apply it on any subset of the qubits involved by simply removing the couplings between certain qubits.

The global entangling gates demand fully connected couplings among all of the involved qubits, which naturally emerge in trapped-ion systems. Ion qubits in a linear chain are entangled by coupling to the collective motional modes, typically through Raman laser beams, as shown in Fig. 1b. Raman beams with beat-note frequencies $\omega_0 \pm \mu$ lead to a qubit-state-dependent force on each qubit site[25]. Here, μ, which has a value around the frequencies of the motional modes, is the detuning from the energy splitting of the qubit, ω_0, as shown in Fig. 1c. The time evolution of the system at time τ can be written as[18]

$$U(\tau) = \exp\left[\sum_{j,m}\beta_{j,m}(\tau)\sigma_x^j - i\sum_{j<j'}\theta_{j,j'}\sigma_x^j\sigma_x^{j'}\right] \quad (2)$$

with $\beta_{j,m}(\tau) = \alpha_{j,m}(\tau)\alpha_m^{\dagger} - \alpha_{j,m}^{*}(\tau)\alpha_m$, where a_m ($\alpha_m^{\dagger}$) is the annihilation (creation) operator of the mth mode, $\alpha_{j,m}$ represents the displacement of the mth motional mode of the jth ion (see Supplementary Information) and $\theta_{j,j'}(\tau)$ is the coupling strength between the jth and jth qubits and has the form

$$\theta_{j,j'}(\tau) = -\sum_m\int_0^{\tau}\mathrm{d}t_2\int_0^{t_2}\mathrm{d}t_1\frac{\eta_{j,m}\eta_{j',m}\Omega_j(t_2)\Omega_{j'}(t_1)}{2}\sin\{(\nu_m-\mu)(t_2-t_1) - [\phi_j(t_2)-\phi_{j'}(t_1)]\} \quad (3)$$

where $\eta_{j,m}$ is the scaled Lamb–Dicke parameter, ν_m is the frequency of the mth motional mode, and $\Omega_j(t)$ and $\phi_j(t)$ are the amplitude and the phase of the carrier Rabi frequency on the jth ion, respectively.

The implementation of global entangling gates would be straightforward if we could only drive the centre-of-mass (COM) mode either in the axial or in the radial direction[10,15,16]. The homogeneous ion–motion couplings of the COM modes, $\eta_{j,1} = \eta_{\mathrm{COM}}$, make all of the coupling strengths uniform as

$$\theta_{j,j'}(\tau) = -\frac{\eta_{\mathrm{COM}}^2\Omega^2\tau}{2(\nu_1-\mu)} \quad (4)$$

by ensuring that $\alpha_{j,1}(\tau) = 0$ at time τ with the conditions $\Omega_j(t) = \Omega$ and $\phi_j(t) = 0$ for all of the ions. However, owing to the bunching of an increasing number of motional modes and their crosstalk when the number of ions increases, we have to dramatically slow down the gate speed to isolate the COM mode[20]. Otherwise, inevitably the rest of the modes are also driven. Either of these effects would decrease the gate fidelity, owing to the limited coherence time or undesired inhomogeneous couplings (see details in Methods), as shown in Fig. 1d. Moreover, the COM modes suffer from more severe electrical noise compared with other modes, and the heating rates increase with the number of ions[26], which would further degrade the gate fidelity.

Owing to the lack of scalability of the single-mode approach, we explore the possibility of finding multi-mode schemes for a scalable global N-qubit entangling gate. To apply the global gate $\mathrm{GE}_N(\Theta)$ in

[1]Center for Quantum Information, Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing, China. [2]MOE Key Laboratory of Fundamental Physical Quantities Measurement & Hubei Key Laboratory of Gravitation and Quantum Physics, PGMF and School of Physics, Huazhong University of Science and Technology, Wuhan, China. [3]These authors contributed equally: Yao Lu, Shuaining Zhang, Kuan Zhang. *e-mail: yao-lu14@mails.tsinghua.edu.cn; kimkihwan@mail.tsinghua.edu.cn

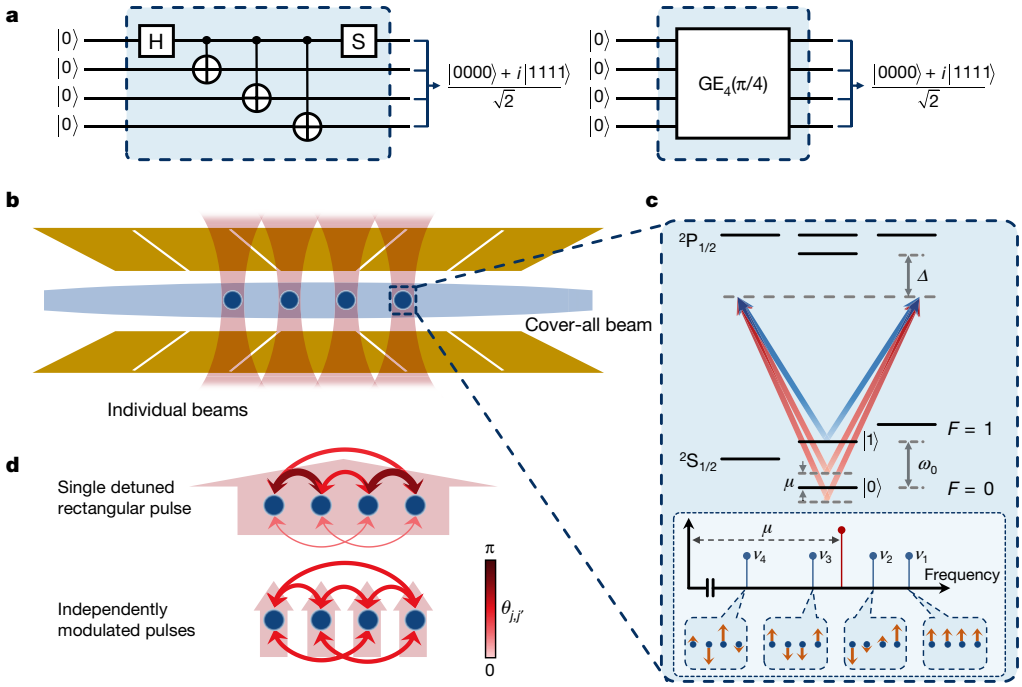


Fig. 1 | Global entangling gate and its experimental implementation. **a**, Efficient construction of a quantum circuit using a global gate. For the generation of the four-qubit GHZ state, we need one Hadamard gate ('H' gate in the figure) and three pairwise entangling gates, which can be replaced by a single global four-qubit entangling gate. The phase gate ('S' gate) at the end of the first circuit is used to compensate for the phase difference between two circuit outputs. **b**, Experimental setup used for the implementation of the global entangling gate. Each ion in the trap encodes a qubit with energy splitting of ω_0, which is individually manipulated by Raman beams: a cover-all beam (blue) and an individual beam addressing a single ion (red). The individually addressed qubits are involved in the global entangling gate. **c**, Energy levels of $^{171}Yb^+$. The Raman beams (with detuning Δ) introduce a qubit-state-dependent force on each ion, with multiple motional modes driven simultaneously at a driving frequency of μ. The patterns of the collective motional modes in the transverse direction and their relative frequencies ν in the spectrum are shown in the inset. **d**, Implementation of the global entangling gate. With a single rectangular pulse, we cannot achieve uniform coupling strengths $\theta_{j,j'}$ on all of the qubit pairs owing to undesired inhomogeneous couplings (see also Methods). Instead, we can achieve uniform coupling by independently modulating the pulses on each ion.

equation (1) using the time evolution of equation (2), we have to close all of the motional trajectories and balance all of the coupling strengths, which lead to the following constraints

$$a_{j,m}(\tau) = 0 \tag{5}$$

$$\theta_{j,j'}(\tau) = \Theta \tag{6}$$

Considering a general situation with N qubits and M collective motional modes, there are $N \times M$ constraints from equation (5) and $\binom{N}{2}$ from equation (6). Therefore, we have to satisfy a total number of $N(N-1)/2 + NM$ constraints. In principle, we can fulfil the constraints by independently modulating the amplitude $\Omega_j(t)$ or the phase $\phi_j(t)$ of the Rabi frequency on each ion in a continuous or a discrete way. In the experimental implementation, we choose discrete phase modulation because we have high-precision controllability on the phase degree of freedom. We divide the total gate operation time into K segments with equal duration and independently modulate the phase on each ion in each segment, which provides $N \times K$ independent variables. Because of the nonlinearity of the constraints, it is challenging to find analytical solutions for the constraints of equations (5) and (6). Therefore, we construct an optimization problem to find numerical solutions. We minimize the objective function of $\sum_{j,m}|\alpha_{j,m}(\tau)|^2$ according to the constraints of equation (6)[21,27,28]. We note that we also use amplitude shaping at the beginning and the end of the operation to minimize fast-oscillating terms due to off-resonant coupling to the carrier transition[29]. Details about the constraints under discrete phase modulation and the construction of the optimization problem are provided in Supplementary Information. Moreover, we note that once we find the solution of the global N-qubit entangling gate, the entangling gate can be applied on any subset of qubits by simply setting $\Omega_j = 0$ for any qubit j outside the subset.

We experimentally implement the global entangling gates in a single linear chain of $^{171}Yb^+$ ions, as shown in Fig. 1b. A single qubit is encoded in the hyperfine levels of the ground-state manifold $^2S_{1/2}$, denoted as $|0\rangle \equiv |F=0, m_F=0\rangle$ and $|1\rangle \equiv |F=1, m_F=0\rangle$ (where F and m_F are the hyperfine and magnetic quantum numbers, respectively), with an energy gap of $\omega_0 = 12.642821$ GHz, as shown in Fig. 1c. The qubits are initialized to state $|0\rangle$ by optical pumping and measured using state-dependent fluorescence detection[30]. The fluorescence is collected by an electron-multiplying charge-coupled device (EMCCD) to realize a site-resolved measurement. After ground-state cooling of the motional modes, coherent manipulations of the qubits are performed by Raman beams produced by a picosecond-pulse laser[31]. One of the Raman beams is broadened to cover all of the ions, whereas the other is divided into several paths that are tightly focused on each ion (referred to as 'individual beam' hereafter). The cover-all beam and the individual beams intersect each other perpendicularly at the ion chain, and drive radial modes mainly along the x direction. Using a multi-channel acousto-optic modulator controlled by a multi-channel arbitrary waveform generator, we realize independent control of the individual beams on each ion, as illustrated in Fig. 1b, similarly to the setup of ref. [14]. Additional information about the experimental setup is provided in Methods.

To test the performance of the global N-qubit entangling gate, we use the $GE_N(\pi/4)$ gate to generate an N-qubit GHZ state and then measure the state fidelity. Starting from the product state $|0\ldots0\rangle$, the GHZ state can be prepared by applying the global entangling gate, while additional single-qubit σ_x rotations by $\pi/2$ are needed if N is odd. After the state preparation, we obtain the state fidelity by measuring the population of the entangled state and the contrast of the parity oscillation[32]. We also use the fidelity of the GHZ state to test the most important

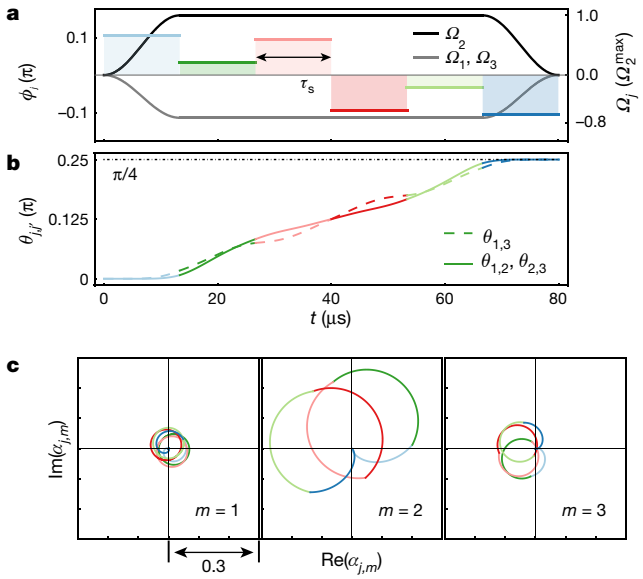


Fig. 2 | Experimental implementation of a global three-qubit entangling gate. a, Pulse scheme with phase and amplitude modulation. The phase ϕ_j is discretely modulated, as shown by the coloured lines. The specific values of the modulated phases are given in Methods. The amplitudes of the Rabi frequencies Ω_j, shown by the black and grey curves, are shaped at the beginning (end) of the gate operation using a sin-squared profile with switching time equal to the duration of a single segment, τ_S. We note that the additional π-phase shift of the middle ion is treated as a negative sign for Ω. **b**, Accumulation of coupling strength $\theta_{j,j'}$ over the evolution time. All of the coupling strengths increase to the desired value of π/4. **c**, Motional trajectories $\alpha_{j,m}$; the first qubit in phase space is shown as an example. Different colours correspond to the different segments in **a**.

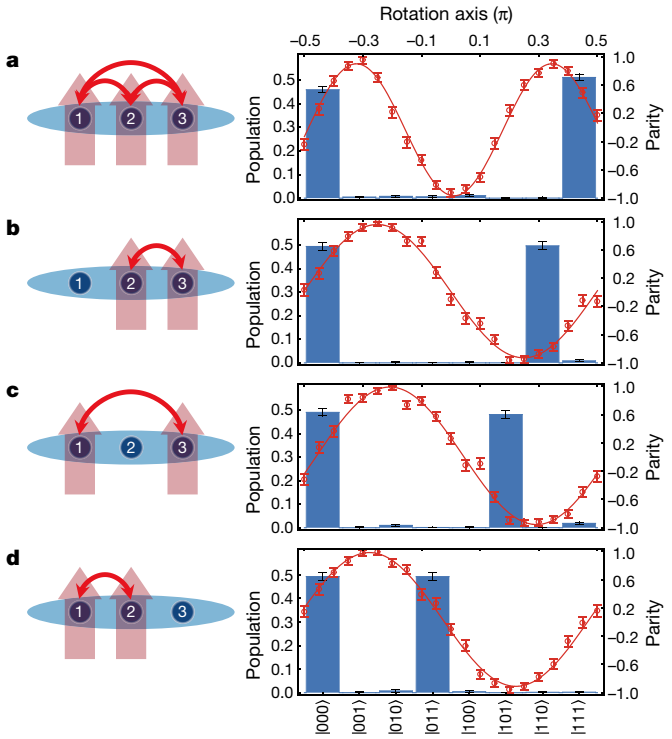


Fig. 3 | Experimental implementation and results of the global entangling gates in three-ion qubits. a–d, The left column shows the operation of the global entangling gate, which can generate entanglement of entire qubits (**a**) or any pair of qubits (**b–d**) by switching on the individual beams on the target ions without changing any modulated patterns. The right column shows the population (blue histogram) and the parity oscillation (red circles, experimental data; red curves, fitting results) of the generated GHZ state. The error bars indicate one standard deviation. **a**, Three-qubit GHZ state with a state fidelity of 95.2% ± 1.5%. **b–d**, Two-qubit GHZ states of qubit pairs (2, 3), (1, 3) and (1, 2), with fidelities of 96.7% ± 1.8%, 97.1% ± 1.9% and 96.5% ± 1.5%, respectively.

feature of the global entangling gate, that is, whether it can be applied on any subset of qubits that are addressed by individual laser beams without changing the modulation pattern.

As a first demonstration of the global entangling gate, we use three $^{171}Yb^+$ ions with the frequencies of the collective motional modes in the x direction $\{\nu_1, \nu_2, \nu_3\} = 2\pi \times \{2.184, 2.127, 2.044\}$ MHz. We choose the detuning μ between the last two modes to be $2\pi \times 2.094$ MHz. The total gate time is fixed at 80 μs and divided into six segments. The details of the phase modulation pattern and the ratio of the amplitude shaping of each ion to the centre one are shown in Fig. 2a. With these parameters, the constraints of equations (5) and (6) are fulfilled, as shown in Fig. 2b, c. We use this global three-qubit entangling gate to prepare the three-qubit GHZ state with a state fidelity of 95.2% ± 1.5% (all uncertainties are one standard deviation), as shown in Fig. 3a. Moreover, by turning off the individual beam on a qubit, we can remove the couplings between that qubit and other qubits, as shown in Fig. 3. In the three-qubit system, the global entangling gates on the subsets become pairwise gates on arbitrary qubit pairs, which are used to generate the two-qubit GHZ states with fidelities higher than 96.5% in the experiment, as shown in Fig. 3b, c.

For a further demonstration of the global entangling gate, we move to a four-qubit system with motional frequencies $\{\nu_1, \nu_2, \nu_3, \nu_4\} = 2\pi \times \{2.186, 2.147, 2.091, 2.020\}$ MHz. The larger system means more constraints, and more segments are required. To realize a global four-qubit entangling gate, we choose the detuning μ to be $2\pi \times 2.104$ MHz and fix the total gate time at 120 μs, which is evenly divided into twelve segments. The pulse scheme is shown in Fig. 4a, b. The number of the constraints in equation (6) increases quadratically with the number of qubits and reaches six in the four-qubit case, as shown in Fig. 4c.

By applying the global four-qubit entangling gate to all of the qubits, we successfully generate a four-qubit GHZ state with a state fidelity of 93.4% ± 2.0%, as shown in Fig. 4d. Similarly, we can prepare a three-qubit GHZ state or a two-qubit GHZ state by only addressing arbitrary three or two qubits, respectively. Experimentally, we choose the qubit set (2, 3, 4) to prepare the three-qubit GHZ state and the qubit pair (1, 3) to prepare the two-qubit GHZ state, with state fidelities of 94.2% ± 1.8% and 95.1% ± 1.6%, respectively, as shown in Fig. 4e, f.

All of the results are corrected to remove detection errors (see Methods). The state fidelities of all of the prepared GHZ states are mainly limited by fluctuations of the tightly focused individual beams and optical-path jittering of the Raman beams (2%–4%). Other infidelity sources in the experiment include drifting of the motional frequencies (1%–2%) and crosstalk of the individual beams with nearby ions (about 1%).

We have presented the experimental realization of global entangling gates, which can increase the efficiency of quantum circuits, using a scalable approach and a trapped-ion platform. The duration of a single global gate is comparable to that of a single pairwise gate with the same total number of ions[20]. Therefore, we clearly observe benefits of the global gates in terms of total gate counts and duration. Moreover, we theoretically optimize the pulse schemes for five and six qubits, and we find that the required number of segments and the gate duration increase linearly with the number of qubits. As long as the solutions to the optimization problem can be determined, we could extend and apply the global entangling gate to a higher number of qubits. Pulse optimization with a large number of qubits is an NP-hard problem, but it could be assisted by a classical machine-learning technique. Furthermore, we can extend the global entangling gate to a general form with arbitrary coupling strengths of $\{\theta_{j,j'}(\tau) = \Theta_{j,j'}\}$, which would further simplify quantum circuits for large-scale quantum computation and simulation[9]. During the preparation of this paper, we became aware of a related study about parallel pairwise entangling gates[33].

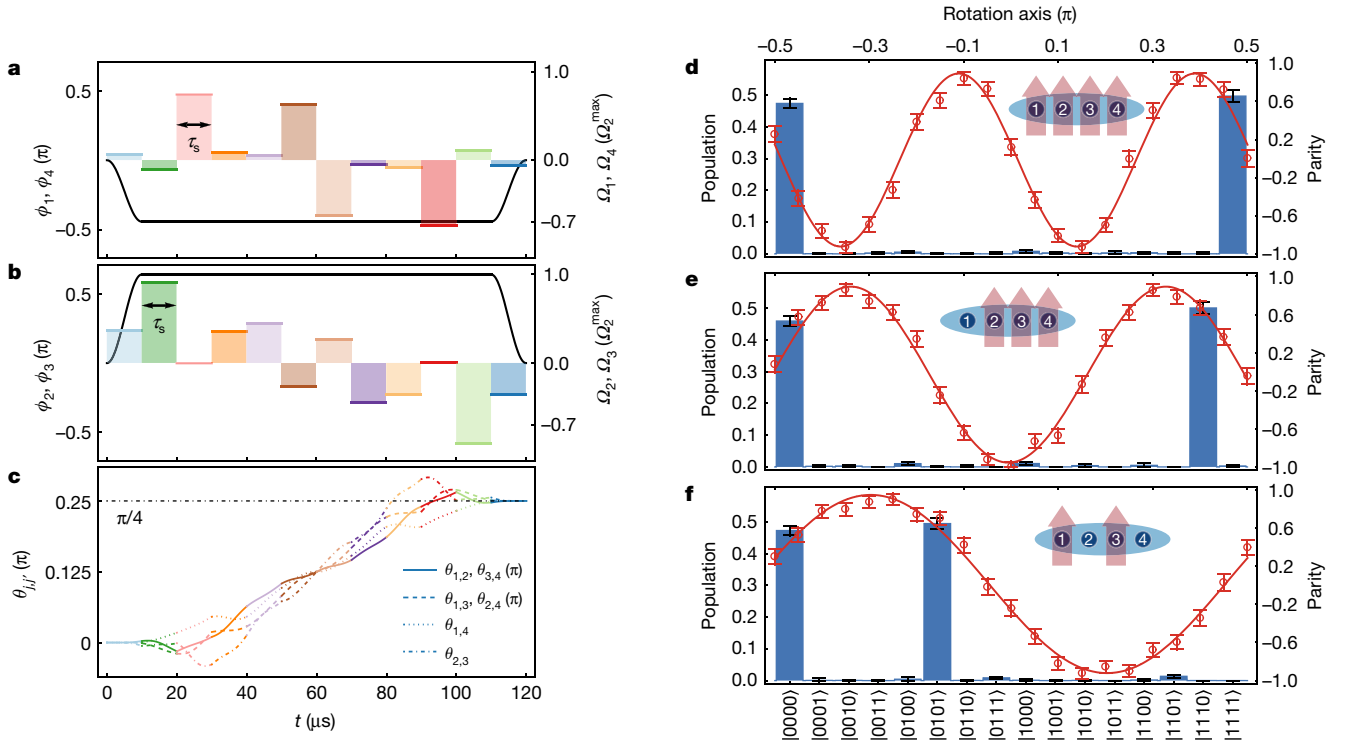


Fig. 4 | Experimental implementation and results of the global entangling gate in a four-ion system. a, b, Pulse scheme with phase and amplitude modulation. Using the symmetry of the system, we set the modulation patterns to be the same for the outer two qubits, (1, 4), and the inner two qubits, (2, 3). The additional π phase shift of each outer ion is treated as a negative sign for the amplitude Ω. The values of the modulated phases and the motional trajectories under this pulse scheme are given in Methods. **c**, Accumulation of coupling strengths $\theta_{j,j'}$ for all of the qubit pairs. The coupling strengths converge to the desired value of π/4 at the end of the gate. **d**–**f**, GHZ states prepared by the global entangling gates. By addressing an arbitrary subset of qubits—for example, (1, 2, 3, 4), (2, 3, 4) and (1, 3)—we can apply the entangling gate on the subset. The frequency of the parity oscillation, which is proportional to the number of addressed qubits, reveals that the prepared state is the GHZ state. Error bars indicate one standard deviation. The state fidelities of the prepared four-, three- and two-qubit GHZ states reach 93.4% ± 2.0%, 94.2% ± 1.8% and 95.1% ± 1.6%, respectively.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41586-019-1428-4.

Received: 15 January 2019; Accepted: 10 May 2019;
Published online 24 July 2019.

1. Shor, P. W. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comput.* **26**, 1484–1509 (1997).
2. Feynman, R. P. Simulating physics with computers. *Int. J. Theor. Phys.* **21**, 467–488 (1982).
3. Lloyd, S. Universal quantum simulators. *Science* **273**, 1073–1078 (1996).
4. Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge Univ. Press, 2010).
5. Casanova, J., Mezzacapo, A., Lamata, L. & Solano, E. Quantum simulation of interacting fermion lattice models in trapped ions. *Phys. Rev. Lett.* **108**, 190502 (2012).
6. Yung, M.-H. et al. From transistor to trapped-ion computers for quantum chemistry. *Sci. Rep.* **4**, 3589 (2015).
7. Ivanov, S. S., Ivanov, P. A. & Vitanov, N. V. Efficient construction of three- and four-qubit quantum gates by global entangling gates. *Phys. Rev. A* **91**, 032311 (2015).
8. Martinez, E. A., Monz, T., Nigg, D., Schindler, P. & Blatt, R. Compiling quantum algorithms for architectures with multi-qubit gates. *New J. Phys.* **18**, 063029 (2016).
9. Maslov, D. & Nam, Y. Use of global interactions in efficient quantum circuit constructions. *New J. Phys.* **20**, 033018 (2018).
10. Kim, K. et al. Entanglement and tunable spin–spin couplings between trapped ions using multiple transverse modes. *Phys. Rev. Lett.* **103**, 120502 (2009).
11. Britton, J. W. et al. Engineered two-dimensional Ising interactions in a trapped-ion quantum simulator with hundreds of spins. *Nature* **484**, 489–492 (2012).
12. Senko, C. et al. Coherent imaging spectroscopy of a quantum many-body spin system. *Science* **345**, 430–433 (2014).
13. Jurcevic, P. et al. Spectroscopy of interacting quasiparticles in trapped ions. *Phys. Rev. Lett.* **115**, 100501 (2015).
14. Debnath, S. et al. Demonstration of a small programmable quantum computer with atomic qubits. *Nature* **536**, 63–66 (2016).
15. Monz, T. et al. 14-qubit entanglement: creation and coherence. *Phys. Rev. Lett.* **106**, 130506 (2011).
16. Lanyon, B. P. et al. Universal digital quantum simulation with trapped ions. *Science* **334**, 57–61 (2011).
17. García-Ripoll, J. J., Zoller, P. & Cirac, J. I. Coherent control of trapped ions using off-resonant lasers. *Phys. Rev. A* **71**, 062309 (2005).
18. Zhu, S.-L., Monroe, C. & Duan, L.-M. Trapped ion quantum computation with transverse phonon modes. *Phys. Rev. Lett.* **97**, 050505 (2006).
19. Steane, A. M., Imreh, G., Home, J. P. & Leibfried, D. Pulsed force sequences for fast phase-insensitive quantum gates in trapped ions. *New J. Phys.* **16**, 053049 (2014).
20. Choi, T. et al. Optimal quantum control of multimode couplings between trapped ion qubits for scalable entanglement. *Phys. Rev. Lett.* **112**, 190502 (2014).
21. Leung, P. H. et al. Robust 2-qubit gates in a linear ion crystal using a frequency-modulated driving force. *Phys. Rev. Lett.* **120**, 020501 (2018).
22. Milne, A. R. et al. Phase-modulated entangling gates robust against static and time-varying errors. Preprint at https://arxiv.org/abs/1808.10462 (2018).
23. Schäfer, V. M. et al. Fast quantum logic gates with trapped-ion qubits. *Nature* **555**, 75–78 (2018).
24. Kaufmann, H. et al. Scalable creation of long-lived multipartite entanglement. *Phys. Rev. Lett.* **119**, 150503 (2017).
25. Haljan, P. C., Brickman, K.-A., Deslauriers, L., Lee, P. J. & Monroe, C. Spin-dependent forces on trapped ions for phase-stable quantum gates and entangled states of spin and motion. *Phys. Rev. Lett.* **94**, 153602 (2005).
26. Lechner, R. et al. Electromagnetically-induced-transparency ground-state cooling of long ion strings. *Phys. Rev. A* **93**, 053401 (2016).
27. Webb, A. E. et al. Resilient entangling gates for trapped ions. *Phys. Rev. Lett.* **121**, 180501 (2018).
28. Shapira, Y., Shaniv, R., Manovitz, T., Akerman, N. & Ozeri, R. Robust entanglement gates for trapped-ion qubits. *Phys. Rev. Lett.* **121**, 180502 (2018).
29. Roos, C. F. Ion trap quantum gates with amplitude-modulated laser beams. *New J. Phys.* **10**, 013002 (2008).
30. Olmschenk, S. et al. Manipulation and detection of a trapped Yb^+ hyperfine qubit. *Phys. Rev. A* **76**, 052314 (2007).
31. Hayes, D. et al. Entanglement of atomic qubits using an optical frequency comb. *Phys. Rev. Lett.* **104**, 140501 (2010).

32. Sackett, C. A. et al. Experimental entanglement of four particles. *Nature* **404**, 256–259 (2000).
33. Figgatt, C. et al. Parallel entangling operations on a universal ion trap quantum computer. *Nature* https://doi.org/10.1038/s41586-019-1427-5 (2019).

Acknowledgements This work was supported by the National Key Research and Development Program of China under grants 2016YFA0301900 and 2016YFA0301901 and the National Natural Science Foundation of China under grants 11574002 and 11504197.

Reviewer information *Nature* thanks Chris Ballance, Roee Ozeri and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Author contributions Y.L., S.Z., K.Z., W.C. and Y.S. developed the experimental system. Y.L. and K.Z., together with J.-N.Z., investigated the theoretical schemes and optimized the pulse sequences. Y.L. and S.Z. obtained the data. K.K. supervised the project. Y.L. led the writing of the manuscript, with contributions from all authors.

Competing interests : The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41586-019-1428-4.
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1428-4.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to Y.L. or K.K.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

METHODS

Comparison of the single-mode and multimode approaches. We compare the single-mode and multimode approaches by numerically calculating the fidelities of GHZ states prepared using these two methods. In the model we only consider the effect of the COM mode and the second mode on a global gate with a radial trap frequency of $2\pi \times 2.18$ MHz and an axial trap frequency varying from $2\pi \times 0.5$ MHz (for three ions) to $2\pi \times 0.32$ MHz (for six ions). These values are consistent with the average experimental spacing of nearby ions of around 4.7 μm. It is difficult to perform a suitable quantum gate with a single axial COM mode at such low axial trap frequencies, owing to high heating rates and poor ground-state cooling, as the gate fidelity would be severely degraded with increasing number of ions. Therefore, we only consider the radial COM mode for the single-mode method.

For the radial COM mode method, we assume that bichromatic fields with detuning μ and time-independent Rabi frequency Ω are applied to all of the ion qubits. To close the trajectories of both modes simultaneously, we let δ_2/δ_1 be an integer r, where $\delta_m = \nu_m - \mu$. Under these assumptions, we can simplify equation (3) to the following form

$$\theta_{j,j'} = -\frac{\delta_1}{|\delta_1|}\frac{\pi(r-1)^2\eta_{\mathrm{COM}}^2\Omega^2}{(\Delta\nu)^2}\left(1+\frac{\eta_{j,2}\eta_{j',2}}{r\eta_{\mathrm{COM}}^2}\right) \tag{7}$$

where $\Delta\nu = |\nu_1 - \nu_2|$ is the frequency difference of the two modes. The gate duration is $\tau = 2\pi|\delta_1|^{-1} = 2\pi|r-1|(\Delta\nu)^{-1}$. An inhomogeneous $\eta_{j,2}$ would imbalance the coupling strengths, as shown in Fig. 1d, for example. We numerically evaluate the fidelities of the created GHZ states by calculating $\mathrm{Fid} = \langle|0\ldots0|\mathrm{GE}_N^{\dagger}(\pi/4)\exp[-i\sum_{j<j'}\theta_{j,j'}\sigma_x^j\sigma_x^{j'}]|0\ldots0\rangle|^2$. The results are summarized in Extended Data Fig. 1. As shown in the figure, to achieve a certain value of state fidelity, the minimal gate duration increases as $N^{2.4}$ with increasing number of ions. We note that we do not include other modes in the simulation, as the inclusion of all modes would lead to further decrease of the fidelity. By contrast, in our multimode approach, we consider the effects of all of the modes. The gate duration increases almost linearly with the number of ions, with unity representing the theoretical fidelity. A shorter gate duration than that of the single-mode approach would suppress the infidelities resulting from the limited coherence time, Raman scattering, motional heating and so on.

Experimental setup. In the experiment the single ion chain is held in a blade trap, in the geometry shown in Extended Data Fig. 2. The average spacing of nearby ions is around 4.7 μm. The Raman beams are produced by a picosecond-pulse laser with a centre wavelength of 377 nm and a repetition rate of about 76 MHz. The ion fluorescence is collected by an objective lens from the top re-entry viewport and then imaged with the EMCCD. The average detection fidelity is 96% for a single ion. The measured population of state, denoted as $P^{\mathrm{meas}} = \{p_{0\ldots0}, \ldots, p_{1\ldots1}\}$, where p_i is the probability of state $|i\rangle$, is calibrated to remove detection errors using the method described in ref. [34], which has been applied to many other experimental demonstrations[12,35]. The matrix of the detection errors (M) is determined experimentally and can be used to reconstruct the real population of the state, $P^{\mathrm{real}} = M^{-1}P^{\mathrm{meas}}$. However, to avoid non-physical results, we utilize the maximum-likelihood method to estimate the real population by minimizing the 2-norm $||P^{\mathrm{meas}} - MP^{\mathrm{real}}||_2$.

Experimental parameters. Here we present the details of the experimental pulse schemes for the global three- and four-qubit entangling gates. The maximal amplitudes of the Rabi frequencies are given using the theoretical Lamb–Dicke parameters

$$\eta_{j,m} = b_{j,m}\frac{2\sqrt{2}\pi}{\lambda}\sqrt{\frac{\hbar}{2M_{\mathrm{Yb}}\nu_m}} \tag{8}$$

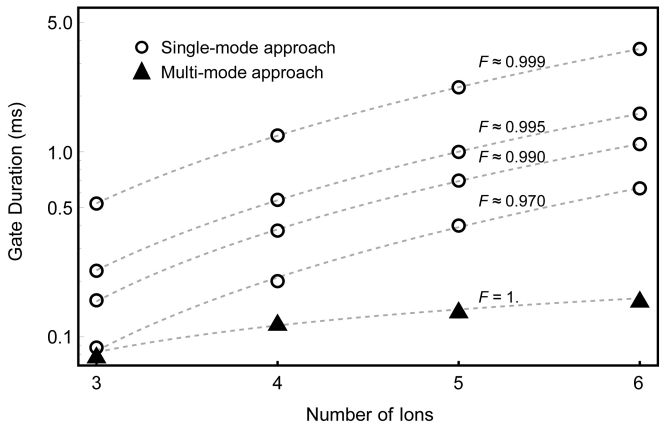
where $b_{j,m}$ is the element of the normal-mode transformation matrix for ion j and motional mode m (ref. [36]), λ is the centre wavelength of the Raman laser, $\hbar$ is the reduced Planck constant and M_{Yb} is the mass of the $^{171}\mathrm{Yb}^+$ ion. For the COM mode, we have typically $\eta \approx 0.08/\sqrt{N}$ for any j in our setup, where N is the number of ions. The values of the modulated phases and amplitudes of the Rabi frequencies obtained from the optimization are shown in Extended Data Tables 1, 2.

In Fig. 2c we show the trajectories of the motional modes in the phase space for the three-qubit situation. In Extended Data Fig. 3, we show the motion trajectories of $\alpha_{j,m}(t)$ for the four-qubit case.

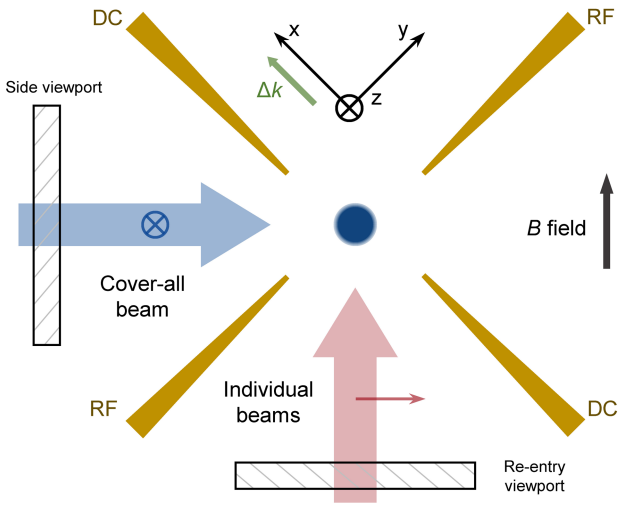
Data availability

All relevant data are available from the corresponding authors upon request.

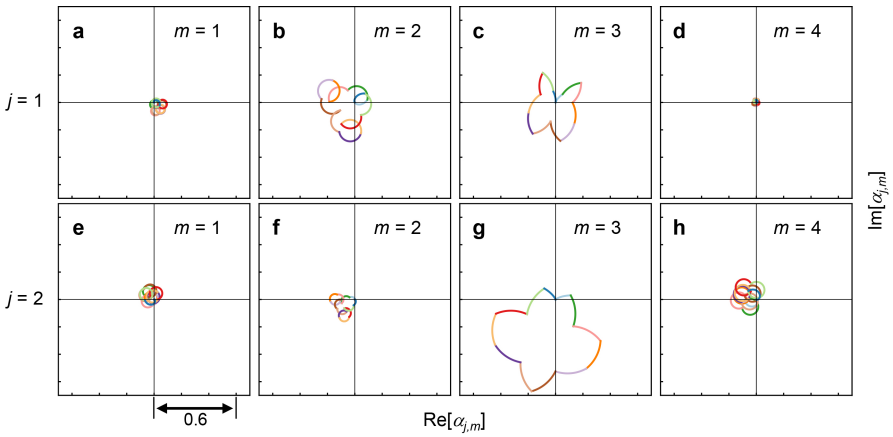
34. Duan, L.-M. & Shen, C. Correcting detection errors in quantum state engineering through data processing. *New J. Phys.* **14**, 1778–1782 (2012).
35. Richerme, P. et al. Non-local propagation of correlations in quantum systems with long-range interactions. *Nature* **511**, 198–201 (2014).
36. James, D. F. V. Quantum dynamics of cold trapped ions with application to quantum computation. *Appl. Phys. B* **66**, 181–190 (1998).



Extended Data Fig. 1 | Comparison between gate durations of single- and multi-mode approaches. For the given trap frequencies, the gate duration τ of the single-mode approach grows faster than linearly ($\tau \approx N^{2.4}$) to maintain the fidelity F when the number of ions, N, increases. The gate duration of the multi-mode approach grows near linearly, with a theoretical fidelity of unity. The vertical axis is on a logarithmic scale.



Extended Data Fig. 2 | Side view of the experimental ion-trap system. The figure shows the structure of the blade trap. The radiofrequency potential is applied to the RF electrodes and the direct-current (DC) electrodes are connected to the direct-current potential. A static magnetic field of $B \approx 6 \times 10^{-4}$ T is applied along the direction shown in the figure. The cover-all beam goes through the side viewport and is focused at the ion-chain position into an elliptical Gaussian beam, with waists of about 30 μm along the ion chain and about 5 μm in the perpendicular direction. The individual beams go through the bottom re-entry viewport and have a focused radius of about 1 μm at the ion position. The average laser power is around 120 mW for the cover-all beam and around 1 mW for each individual beam. The effective wave vector Δk of the two Raman beams is almost in the x direction, and the beams are polarized linearly, perpendicular to each other.



Extended Data Fig. 3 | Motional trajectories in phase space for the global four-qubit entangling gate. Because we apply different modulated-phase patterns to the qubits (1, 4) and (2, 3), the shapes of the motional trajectories in **a**–**d** and **e**–**h** are different.

Extended Data Table 1 | Pulse scheme for the global three-qubit entangling gate

Qubit j		1	2	3
Ω_j^{max} (MHz)		$-2\pi \times 0.181$	$2\pi \times 0.253$	$-2\pi \times 0.181$
$\phi_{j,k}$ (π)	1	0.104	0.104	0.104
	2	0.033	0.033	0.033
	3	0.095	0.095	0.095
	4	-0.095	-0.095	-0.095
	5	-0.033	-0.033	-0.033
	6	-0.104	-0.104	-0.104

Here, Ω_j^{max} refers to the maximal amplitude of the Rabi frequency on the jth qubit during pulse shaping and $\phi_{j,k}$ refers to the value of the modulated phase on the jth qubit in the kth segment.

Extended Data Table 2 | Pulse scheme for the global four-qubit entangling gate

Qubit j		1	2	3	4
Ω_j^{max} (MHz)		$-2\pi \times 0.117$	$2\pi \times 0.168$	$2\pi \times 0.168$	$-2\pi \times 0.117$
$\phi_{j,k}$ (π)	1	0.041	0.231	0.231	0.041
	2	-0.070	0.579	0.579	-0.070
	3	0.472	-0.001	-0.001	0.472
	4	0.054	0.230	0.230	0.054
	5	0.035	0.285	0.285	0.035
	6	0.402	-0.170	-0.170	0.402
	7	-0.402	0.170	0.170	-0.402
	8	-0.035	-0.285	-0.285	-0.035
	9	-0.054	-0.230	-0.230	-0.054
	10	-0.472	0.001	0.001	-0.472
	11	0.070	-0.579	-0.579	0.070
	12	-0.041	-0.231	-0.231	-0.041

The definitions of Ω_j^{max} and $\phi_{j,k}$ are as in Extended Data Table 1.

LETTER

https://doi.org/10.1038/s41586-019-1427-5

Parallel entangling operations on a universal ion-trap quantum computer

C. Figgatt[1,2,3,6]*, A. Ostrander[2,3], N. M. Linke[1,2], K. A. Landsman[1,2,3], D. Zhu[1,2,3], D. Maslov[1,2,3,4,7] & C. Monroe[1,2,3,5]

The circuit model of a quantum computer consists of sequences of gate operations between quantum bits (qubits), drawn from a universal family of discrete operations[1]. The ability to execute parallel entangling quantum gates offers efficiency gains in numerous quantum circuits[2–4], as well as for entire algorithms—such as Shor's factoring algorithm[5]—and quantum simulations[6,7]. In circuits such as full adders and multiple-control Toffoli gates, parallelism can provide an exponential improvement in overall execution time through the divide-and-conquer technique[8]. More importantly, quantum gate parallelism is essential for fault-tolerant error correction of qubits that suffer from idle errors[9,10]. However, the implementation of parallel quantum gates is complicated by potential crosstalk, especially between qubits that are fully connected by a common-mode bus, such as in Coulomb-coupled trapped atomic ions[11,12] or cavity-coupled superconducting transmons[13]. Here we present experimental results for parallel two-qubit entangling gates in an array of fully connected trapped $^{171}Yb^{+}$ ion qubits. We perform a one-bit full-addition operation on a quantum computer using a depth-four quantum circuit[4,14,15], where circuit depth denotes the number of runtime steps required. Our method exploits the power of highly connected qubit systems using classical control techniques and will help to speed up quantum circuits and achieve fault tolerance in trapped-ion quantum computers.

Trapped atomic ions are among the most advanced qubit platforms[11,12], with atomic-clock precision and the ability to perform gate operations in a fully connected and reconfigurable qubit network[16]. The high connectivity between trapped-ion qubits[17] is mediated by optical forces applied to their collective motion[18], and can be scaled in a modular fashion using a variety of methods[11,12]. Although the all-to-all interactions provided by ion-trap systems are powerful tools that can be used to create large global entangled states and perform large analogue quantum simulations[19–21], they also present substantial, previously unaddressed challenges for implementing the full control necessary for independent, parallel entangling operations. Additionally, although previous efforts have demonstrated the control necessary for individual addressing and universal gate sets[16,22], concurrent, arbitrary control of individual ions—which is necessary to enact parallel operations—had not previously been demonstrated. We note that global operations cannot perform different operations on different ions at the same time; symmetry-breaking control is required. Within a single large chain of ions, gates can be realized by appropriately shaping the laser pulses that drive selected ions within the chain. Here, the target qubits become entangled through their Coulomb-coupled motion, and the laser pulse is modulated so that the motional modes are disentangled from the qubits at the end of the operation[23–25]. The execution of multiple parallel gates in this way requires more complex pulse shapes, not only to disentangle the motion but also to entangle exclusively the intended qubit pairs. We achieve this type of parallel operation by designing appropriate optical pulses using nonlinear optimization techniques.

We perform parallel gate operations on a chain of five atomic $^{171}Yb^{+}$ ions, using resonant laser radiation to laser-cool, initialize and measure the qubits. Coherent quantum gate operations are achieved by applying counterpropagating Raman beams from a single mode-locked laser, which form beat notes near the qubit difference frequency. Single-qubit gates are generated by tuning the Raman beat note to the qubit frequency splitting, ω_0, and driving resonant Rabi rotations (R gates) of defined phase and duration. Two-qubit (XX) gates are realized by illuminating two ions with beams that have beat-note frequencies near the motional sidebands, creating an effective Ising interaction between the ions via transient entanglement through the modes of motion[11,12,18]. We use an amplitude-modulated pulse-shaping scheme that provides high-fidelity entangling gates on any ion pair[16,24,25]; frequency[26] or phase[27] modulation of the laser pulses would also suffice. (See Methods for additional experimental details.) A related method was developed in parallel to ours to create multi-qubit entangled states in ion chains[28].

To perform parallel entangling operations involving M independent pairs of qubits in a chain of $N \geq 2M$ ions with N motional modes at frequencies ω_k, a shaped qubit-state-dependent force is applied to the ions involved using bichromatic beat notes at $\omega_0 \pm \mu$, resulting in the evolution operator[23,24,29]

$$U_{||}(\tau) = \exp\left[\sum_{i=0}^{2M} \hat{\phi}_i(\tau)\sigma_i^x + i\sum_{i<j}^{2M} \chi_{ij}(\tau)\sigma_i^x\sigma_j^x\right] \quad (1)$$

where τ is the gate time and $\sigma_{i,j}^x$ is the Pauli spin matrix for qubit i. The first operator describes state-dependent displacements of each mode k in phase space[24,29], with $\hat{\phi}_i(\tau) = \sum_k [\alpha_{i,k}(\tau)\hat{a}_k^\dagger - \alpha_{i,k}^*(\tau)\hat{a}_k]$ and accumulated displacement value

$$\alpha_{i,k}(\tau) = \int_0^\tau \eta_{i,k}\Omega_i(t)\sin(\mu t)e^{i\omega_k t}dt \quad (2)$$

Here, $\hat{a}_k^\dagger$ and $\hat{a}_k$ are the raising and lowering operators for mode k, $\eta_{i,k}$ is the Lamb–Dicke parameter coupling qubit i to mode k, and $\Omega_i(t)$ is the Rabi frequency of the ith ion, which is proportional to the amplitude-modulated laser intensity applied on the ion. To generate independent XX gates, we implement separate control signals for each of the M ion pairs that we want to entangle, thereby providing enough parameters to simultaneously entangle only the desired ion pairs. The parameter χ_{ij} in equation (1) entangles qubits i and j and is given by

$$\chi_{ij}(\tau) = 2\int_0^\tau dt' \int_0^{t'} dt \sum_k \eta_{i,k}\eta_{j,k}\Omega_i(t)\Omega_j(t)\sin(\mu t)\sin(\mu t') \sin[\omega_k(t'-t)] \quad (3)$$

At the end of the gate operation, the $2MN$ accumulated displacement values in equation (2) (for the $2M$ ions involved and for N modes) should vanish so that all mode trajectories close in phase space and

[1]Joint Quantum Institute, University of Maryland, College Park, MD, USA. [2]Department of Physics, University of Maryland, College Park, MD, USA. [3]Joint Center for Quantum Information and Computer Science, University of Maryland, College Park, MD, USA. [4]National Science Foundation, Alexandria, VA, USA. [5]IonQ Inc., College Park, MD, USA. [6]Present address: Honeywell, Broomfield, CO, USA. [7]Present address: IBM Thomas J. Watson Research Center, Yorktown Heights, NY, USA. *e-mail: cfiggatt@umd.edu

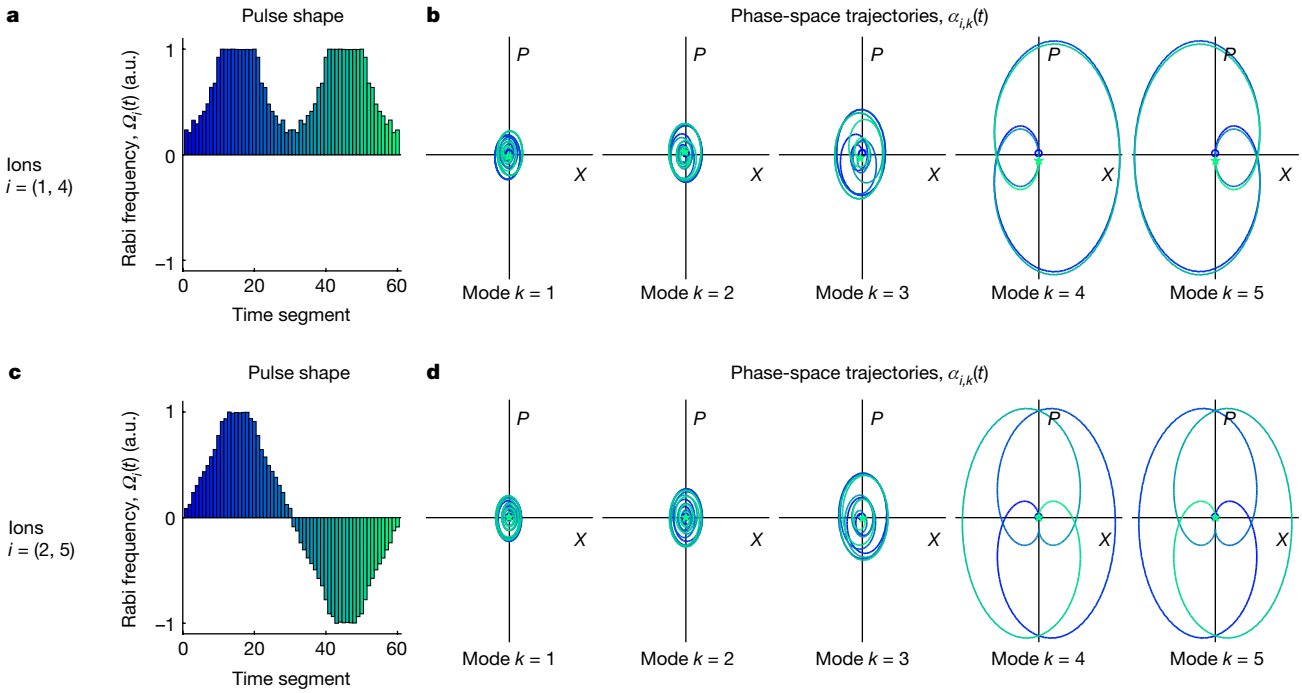


Fig. 1 | Parallel-gate pulse solutions. a–d, Laser pulse shape solutions (**a**, **c**) and theoretical phase-space trajectories $\alpha_{i,k}$ for each mode k correlated with ion i (**b**, **d**) for parallel XX gates on ions (1, 4) (**a**, **b**) and ions (2, 5) (**c**, **d**). The pulse shape solutions are expressed in terms of the time-dependent Rabi frequency $\Omega_i(t)$ experienced by both ions in each pair and is broken into $S = 60$ segments with a total gate time of 250 μs. Negative Rabi frequencies correspond to an inverted phase of the beat note. The five modes of motion have frequencies $\omega_k/2\pi = \{3.045, 3.027, 3.005, 2.978, 2.946\}$ MHz, and with a constant laser beat-note detuning of $\mu = 2.962$ MHz, the nearby modes 4 and 5 experience the largest displacements. The phase-space trajectories in **b**, **d** begin at the blue circles and follow continuous paths to the green stars, with the colour shading of the trajectory corresponding to the pulse shape in time in **a**, **c**. The sum of the normalized area enclosed by all five modes is set to $\pi/4$. X and P designate position and momentum, respectively. a.u., arbitrary units.

there is no residual qubit-motion entanglement. For each of the M desired entangled pairs, we require $\chi_{ij} = \pi/4$ for maximal entanglement (or other non-zero values for partial entanglement); for the other pairs of qubits, whose interactions represent crosstalk, $\chi_{ij} = 0$. This yields a total of $2MN + \binom{2M}{2} = 2MN + \frac{(2M)!}{2!(2M-2)!}$ constraints for designing appropriate pulse sequences $\Omega_i(t)$ to implement M parallel entangling gates. To provide optimal control during the gate and fulfill these constraints, we divide the laser pulse at ion i into S segments of equal time duration τ/S and vary the amplitude in each segment as an independent variable.

Whereas the $2MN$ motional mode constraints (equation (2)) are linear with respect to the control parameters $\Omega_i(t)$, the $\binom{2M}{2}$ entanglement constraints (equation (3)) are quadratic. Finding pulse solutions to this non-convex quadratically constrained quadratic program is an NP-hard problem in general. Because analytical approaches are intractable, we use numerical optimization techniques to find solutions. Further discussion of the constraint problem setup and derivation of the fidelity of simultaneous XX gate operations as a function of the above control parameters is provided in Supplementary Information and ref. [30].

Parallel gates are designed for two independent ion pairs in a five-ion chain. Pulse sequences are designed by solving an optimization problem that takes into account the laser power and the constraints on parameters α and χ (see Supplementary Information). Sequences are calculated for a gate time of $\tau_{gate} = 250$ μs, which is comparable to the standard two-qubit XX gates already used on the experiment, as described in ref. [16], and for a range of detunings μ. This generates a selection of solutions, which are tested on the experimental setup; the solution generating the highest-quality gate using the lowest amount of power is chosen.

Experimental gates are found for six ion-pair combinations: {(1, 4), (2, 5)}; {(1, 2), (3, 4)}; {(1, 5), (2, 4)}; {(1, 4), (2, 3)}; {(1, 3), (2, 5)} and {(1, 2), (4, 5)}. Figure 1 shows the pulse sequence applied to each entangled pair to construct a set of parallel two-qubit gates on ions (1, 4) and (2, 5), as well as the trajectories of each mode–pair interaction in phase space. The five transverse motional modes in this five-ion chain have sideband frequencies $\{\omega_k/2\pi\} = \{3.045, 3.027, 3.005, 2.978, 2.946\}$ MHz, where mode 1 is the common mode at 3.045 MHz. The phase-space trajectories show that modes 4 and 5, which are closest to the selected detuning of $\mu = 2.962$ MHz, exhibit the greatest displacement and contribute the most to the final spin–spin entanglement by enclosing a larger area of phase space. Negative-amplitude pulses are implemented by applying a phase shift of π to the control signal, allowing the entangling pairs to continue accumulating entanglement while cancelling out accumulated entanglement with crosstalk pairs. Consequently, all of the pulse solutions feature similar patterns with symmetric phase flips on one pair to cancel out crosstalk entanglement. Pulse shapes and phase-space trajectories for additional solutions are given in ref. [30].

We characterize the experimental gate fidelities by measuring the selected output qubits in different bases and extracting the parity as a witness operator[31], as described in Supplementary Information. Fitted parity curves are shown in Fig. 2. Entangling-gate fidelities are typically 96%–99%, with crosstalk errors of a few per cent. Crosstalk fidelities are estimated by fitting the crosstalk-pair populations and parity in the same way as above. A fidelity of 25% indicates a complete statistical mixture, which all of the pairs are close to; any fidelity above that value represents an unwanted correlation or a small amount of entanglement, and this difference is reported here as the crosstalk error. The uncertainties given are statistical. All data have been corrected for state-preparation and measurement errors of 3%–5%, as described in refs [16,30].

As an example application of a parallel operation that is useful for error-correction codes[3], we apply a pair of controlled NOT (CNOT) gates in parallel on two pairs of ions. The CNOT gate sequence (a compiled version with R and XX gates is presented in ref. [16]) is performed

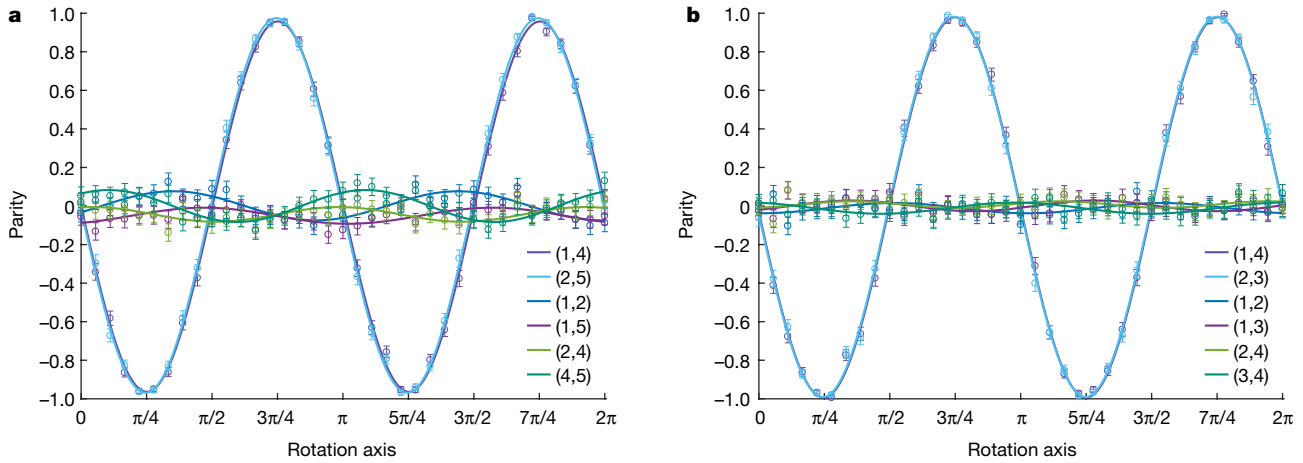


Fig. 2 | Experimental gate fidelities for parallel two-qubit entangling gates. a, **b**, Parity curves used to calculate fidelities for parallel XX gates on two example sets of ions. Circles indicate data and matching-colour lines represent calculated fits. The key specifies the ion pair corresponding to each parity curve, including the two gate ion pairs (the first two ion pairs in the key) and the four crosstalk ion pairs. Additional data are given in Methods. **a**, Ions (1, 4) and (2, 5) yield fidelities of 96.5(4)% and 97.8(3)%, respectively, for the corresponding entangled pairs, with an average crosstalk error of 3.6(3)%. **b**, Ions (1, 4) and (2, 3) yield fidelities of 98.8(3)% and 99.0(3)%, respectively, for the corresponding entangled pairs, with an average crosstalk error of 1.4(3)%. The quoted errors are statistical (1 s.d.).

simultaneously on the pair (1, 4), with ion 1 acting as the control and ion 4 acting as the target, and on the pair (2, 3), with ion 2 acting as the control and ion 3 acting as the target. The simultaneous CNOT gates are applied for each of the 16 possible bitwise inputs, and population data for the 16 possible bitwise outputs, with an average process fidelity of 94.5(2)%, are shown in Fig. 3. All uncertainties correspond to one standard deviation.

Another application that benefits from the use of parallel entangling operations is the quantum full adder. In modern classical computing, a full adder is a basic circuit that can be cascaded to add many-bit numbers, which can be found in processors as a component of arithmetic logic units or performing low-level operations such as computing register addresses. In quantum computing, adders can be used in a similar fashion to perform arithmetic operations over quantum registers (for example, ref. [6]); some algorithms are dominated by adders—notably, Shor's integer factoring algorithm. The quantum full adder requires four qubits: three for the primary inputs x, y and the carry bit C_{in}, and the fourth initialized to $|0\rangle$. The four outputs consist of: the first input, x, simply continuing through; y', which carries $x \oplus y$ (an additional CNOT operation can be added to extract y if desired), where $\oplus$ denotes bitwise addition modulo 2, or XOR; and the sum S and output carry bit C_{out}, which together comprise the two-bit result of summing x, y and C_{in}, where C_{out} is the most significant bit—and hence becomes the carry bit to the next adder in a cascade—and S is the least significant bit. We can also write the sum as $S = x \oplus y \oplus C_{\text{in}}$ and the output carry as $C_{\text{out}} = (x \cdot y) \oplus (C_{\text{in}} \cdot (x \oplus y))$, where $\cdot$ denotes bitwise multiplication, or AND. Feynman first designed such a circuit using CNOT and Toffoli gates[14] (Fig. 4a), which would require 12 XX gates to implement on an ion-trap quantum computer. A more efficient circuit requires at most six two-qubit interactions[4] and features a gate depth of only 4 if simultaneous two-qubit operations are available, as shown by the dashed outlines in Fig. 4b.

The full adder is implemented using two different parallel XX gate configurations, as well as the single-qubit rotations and additional XX gates shown in Extended Data Fig. 4. The parallel gates, a CNOT and its square root (see Methods), require different amounts of entanglement, equivalent to implementing a fully entangling $\text{XX}(\chi_{ij} = \pi/4)$ gate and a partially entangling $\text{XX}(\chi_{ij} = \pi/8)$ gate in parallel. This is experimentally implemented by adjusting the optical power supplied to each gate independently; a discussion of the calibration independence of these parallel gates and fidelity data for such an operation are given in Methods. The inputs x, y, C_{in} and 0 are mapped to the qubits 1, 2, 4 and 5, respectively. Figure 4c shows the data resulting from implementing this computation, with all eight possible bitwise inputs on the three input qubits, and displays the populations in all of the 16 possible bitwise outputs on the four qubits used. The data yield an average process fidelity of 83.3(3)%.

Faster serial two-qubit gates can be accomplished with more optical power, but this speedup is limited by sideband resolution, and this limitation gets worse as the processor size grows owing to spectral crowding. Parallel two-qubit operations are a tool to speed up computation that avoids this problem. This work presents parallel operations with gate times comparable to that of simple two-qubit gates in the same system; tradeoffs between optical intensity and gate time are discussed in Methods. The control scheme presented here for parallel two-qubit entangling gates in ions also suggests a method for performing multi-qubit entanglement in a single operation, which is discussed in Supplementary Information.

When pre-calculating optimal solutions, the number of constraints grows polynomially with the number of ions and entangling pairs. Two parallel XX gates in a chain of N ions require $4N + 6 = O(N)$ constraints, so the problem size grows linearly with N. Entangling more pairs in parallel enlarges the problem size quadratically: entangling M pairs involves the interactions of $2M$ ions, yielding

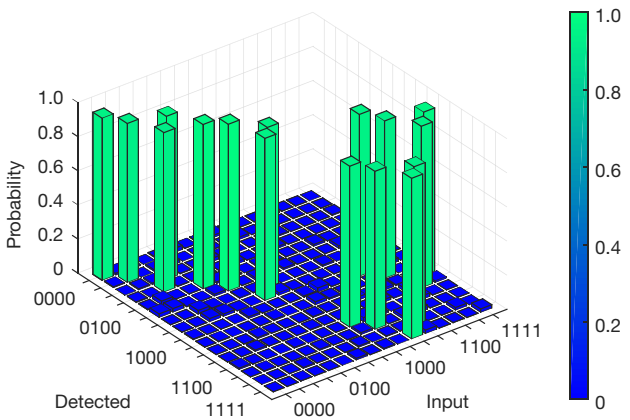


Fig. 3 | Experimental data for parallel CNOT gates. Data for simultaneous CNOT gates on ions (1, 4) and (2, 3), with an average process fidelity of 94.5(2)%. All possible binary input states are tested, and the probability of detecting each possible output state is shown for each input state. The quoted errors are 1 s.d.

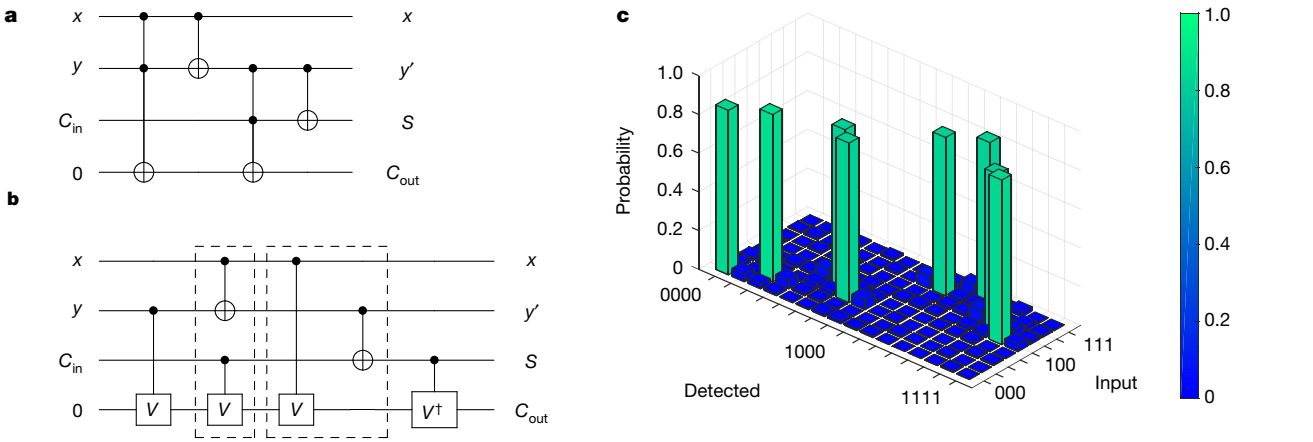


Fig. 4 | Quantum full adder. a, The original quantum full-adder circuit proposed by Feynman in 1985[14], with a two-qubit gate depth of 12. **b**, Optimized full adder with a two-qubit gate depth of 4 (ref. [4]). The two parallel two-qubit operations are outlined in dashed boxes. The $C(V)$ and $C(V^{\dagger})$ (where $V=\sqrt{\mathrm{NOT}}$) operations are the square root of the CNOT gate and its complex conjugate, respectively (see Methods) The circuits in **a** and **b** use standard quantum circuit notation, where each horizontal line denotes a single qubit, labelled at the input and output, and connecting vertical lines depict multi-qubit interactions, including CNOT gates (dot on the control qubit, ⊕ on the target qubit), Toffoli gates (dots on two control qubits, ⊕ on the target qubit) and controlled unitary gates (dot on the control qubit, unitary name on the target qubit.) **c**, Data for the experimental implementation of the full adder using simultaneous two-qubit gates on ions (1, 2, 4, 5), with an average process fidelity of 83.3(3)%. All of the eight possible bitwise input states on the three input qubits are tested, and the probability of detecting each possible output state on the four output qubits is shown for each input state. The quoted errors are 1 s.d.

$\binom{2M}{2}=2M^2-M=O(M^2)$ spin–spin interactions to control and $2MN$ spin–motional entanglements to close. Scaling both the number of entangled pairs M and the number of ions N in the chain therefore gives a total number of constraints of $2MN+2M^2-M=O(M^2+MN)$. On very long chains, not all ion–ion connections will be directly available[32], reducing the number of quadratic constraints on crosstalk pairs that must be considered and thus setting an upper bound on the scaling. Furthermore, when a set of parallel quantum gates is applied on target ions that are m atomic positions apart in a long chain, the effective crosstalk errors fall off[33] as $(1/m)^3$. This implies an ability to perform parallel gate operations in separate local zones in a long chain with little pulse-complexity overhead or fidelity loss.

Several lines of future inquiry may help increase the theoretical solution fidelity. Easing constraints on the power needed may enable the calculation of higher-fidelity solutions, although increasing the power in the experiment can exacerbate errors that arise from noise on the Raman beam. Investigating whether the constraint matrices in equation (11) of Supplementary Information can be modified to become positive or negative semidefinite may provide improvements, as convex quadratically constrained quadratic programs are readily solved using semidefinite programming techniques, and could enable higher-fidelity solutions. However, these are all problems of overhead. Once a high-quality gate solution is implemented in the experiment, no further calculations are needed; only a single calibration is required to compensate for Rabi frequency drifts.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41586-019-1427-5.

Received: 26 November 2018; Accepted: 3 June 2019;
Published online 24 July 2019.

1. Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge Univ. Press, 2011).
2. Cleve, R. & Watrous, J. Fast parallel circuits for the quantum Fourier transform. In *Proc. 41st Annual Symposium on Foundations of Computer Science* 526–536 (IEEE, 2000).
3. Maslov, D. Linear depth stabilizer and quantum Fourier transformation circuits with no auxiliary qubits in finite neighbor quantum architectures. *Phys. Rev. A* **76**, 052310 (2007).
4. Maslov, D., Dueck, G. W., Miller, D. M. & Negrevergne, C. Quantum circuit simplification and level compaction. *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.* **27**, 436–444 (2008).
5. Fowler, A. G., Devitt, S. J. & Hollenberg, L. C. L. Implementation of Shor's algorithm on a linear nearest neighbour qubit array. *Quantum Inf. Comput.* **4**, 237–251 (2004).
6. Nam, Y. & Maslov, D. Low-cost quantum circuits for classically intractable instances of the Hamiltonian dynamics simulation problem. *npj Quantum Inf.* **5**, 44 (2019).
7. Kivlichan, I. D. et al. Quantum simulation of electronic structure with linear depth and connectivity. *Phys. Rev. Lett.* **120**, 110501 (2018).
8. Cormen, T. H., Leiserson, C. E., Rivest, R. L. & Stein, C. *Introduction to Algorithms* 3rd edn (MIT Press, 2009).
9. Steane, A. M. Space, time, parallelism and noise requirements for reliable quantum computing. *Fortschr. Phys.* **46**, 443–457 (1999).
10. Aharonov, D. & Ben-Or, M. Fault-tolerant quantum computation with constant error rate. *SIAM J. Comput.* **38**, 1207–1282 (2008).
11. Wineland, D. & Blatt, R. Entangled states of trapped atomic ions. *Nature* **453**, 1008–1014 (2008).
12. Monroe, C. & Kim, J. Scaling the ion trap quantum processor. *Science* **339**, 1164–1169 (2013).
13. Devoret, M. H. & Schoelkopf, R. J. Superconducting circuits for quantum information: an outlook. *Science* **339**, 1169–1174 (2013).
14. Feynman, R. P. Quantum mechanical computers. *Optics News* **11**, 11–20 (1985).
15. Draper, T. G., Kutin, S. A., Rains, E. M. & Svore, K. M. A logarithmic-depth quantum carry-lookahead adder. *Quantum Inf. Comput.* **6**, 351–369 (2006).
16. Debnath, S. et al. Demonstration of a small programmable quantum computer with atomic qubits. *Nature* **536**, 63–66 (2016).
17. Linke, N. M. et al. Experimental comparison of two quantum computing architectures. *Proc. Natl Acad. Sci. USA* **114**, 3305–3310 (2017).
18. Mølmer, K. & Sørensen, A. Multiparticle entanglement of hot trapped ions. *Phys. Rev. Lett.* **82**, 1835–1838 (1999).
19. Islam, R. et al. Emergence and frustration of magnetism with variable-range interactions in a quantum simulator. *Science* **340**, 583–587 (2013).
20. Bohnet, J. G. et al. Quantum spin dynamics and entanglement generation with hundreds of trapped ions. *Science* **352**, 1297–1301 (2016).
21. Kokail, C. et al. Self-verifying variational quantum simulation of lattice models. *Nature* **569**, 355–360 (2019).
22. Friis, N. et al. Observation of entangled states of a fully-controlled 20 qubit system. *Phys. Rev. X* **8**, 021012 (2018).
23. Zhu, S.-L., Monroe, C. & Duan, L.-M. Trapped ion quantum computation with transverse phonon modes. *Phys. Rev. Lett.* **97**, 050505 (2006).
24. Zhu, S.-L., Monroe, C. & Duan, L.-M. Arbitrary-speed quantum gates within large ion crystals through minimum control of laser beams. *Europhys. Lett.* **73**, 485–491 (2006).
25. Choi, T. et al. Optimal quantum control of multimode couplings between trapped ion qubits for scalable entanglement. *Phys. Rev. Lett.* **112**, 190502 (2014).
26. Leung, P. H. et al. Robust 2-qubit gates in a linear ion crystal using a frequency-modulated driving force. *Phys. Rev. Lett.* **120**, 020501 (2018).

27. Green, T. J. & Biercuk, M. J. Phase-modulated decoupling and error suppression in qubit-oscillator systems. *Phys. Rev. Lett.* **114**, 120502 (2015).
28. Lu, Y. et al. Global entangling gates on arbitrary ion qubits. *Nature* https://doi.org/10.1038/s41586-019-1428-4 (2019).
29. García-Ripoll, J. J., Zoller, P. & Cirac, J. I. Trapped ion quantum computation with transverse phonon modes. *Phys. Rev. A* **71**, 062309 (2005).
30. Figgatt, C. *Building and Programming a Universal Ion Trap Quantum Computer*. PhD thesis, Univ. of Maryland (2018).
31. Sackett, C. A. et al. Experimental entanglement of four particles. *Nature* **404**, 256–259 (2000).
32. Lin, G.-D. et al. Large-scale quantum computation in an anharmonic linear ion trap. *Europhys. Lett.* **86**, 60004 (2009).
33. Landsman, K. A. et al. Two-qubit entangling gates within arbitrarily long chains of trapped ions. Preprint at https://arxiv.org/abs/1905.10421 (2019).

Acknowledgements We thank S.-T. Wang, Z. Gong, S. Debnath, P. H. Leung, Y. Wu and L. Duan for discussions. Circuits were drawn using the qcircuit.tex package. This work was supported by the ARO with funds from the IARPA LogiQ programme, the AFOSR MURI programme and the NSF Physics Frontier Center at JQI. This material is partially based on work supported by the National Science Foundation during D.M.'s assignment at the Foundation. Any opinion, finding, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Author contributions C.F., A.O., N.M.L., K.A.L., D.M. and C.M. designed the research; C.F., N.M.L., K.A.L., D.Z. and C.M. collected and analysed the data; C.F. performed the theory derivations; A.O. performed the pulse sequence optimizations; C.F., A.O., N.M.L., K.A.L., D.Z., D.M. and C.M. contributed to the manuscript.

Competing interests C.M. is co-founder and Chief Scientist at IonQ, Inc.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41586-019-1427-5.
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1427-5.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to C.F.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

METHODS

Experimental setup and error sources. The experiments are performed on a linear chain of five trapped $^{171}Yb^{+}$ ions that are laser-cooled to near their ground state. We designate the qubit as the $|0\rangle \equiv |F=0, m_F=0\rangle$ and $|1\rangle \equiv |F=1, m_F=0\rangle$ hyperfine-split electronic states of the ion's $^2S_{1/2}$ manifold[34], which are first-order magnetic-field-insensitive clock states with a splitting of 12.642821 GHz (F and m_F are the hyperfine and magnetic quantum numbers, respectively). Coherent operations are performed by counterpropagating Raman beams from a single 355-nm mode-locked laser. Spontaneous photon scattering errors are very small in our system (probability of $<10^{-4}$ during a gate) owing to the large detuning of the Raman beams (33 and 67 THz) from the resonant S–P transitions. The first Raman beam is a global beam applied to the entire chain, and the second one is split into individual addressing beams to target each ion qubit[16]. Additionally, a multi-channel arbitrary waveform generator provides separate radiofrequency control signals to each ion's individual addressing beam, providing the individual phase, frequency and amplitude controls that are necessary to execute independent two-qubit operations in parallel. Qubits are initialized to the $|0\rangle$ state using optical pumping and are read out by separate channels of a multi-channel photomultiplier tube array using state-dependent fluorescence.

Measured parallel-gate and algorithmic-process fidelities are reduced from the theoretically calculated fidelities primarily due to engineering imperfections in the experimental system. Beam-pointing instabilities of the individual Raman beams cause Rabi frequency fluctuations, which produce small random coherent errors during gates and comprise the predominant source of error in the system. Crosstalk between individual ion-addressing Raman beams and imperfect compensation of inhomogeneous Stark shifts across the ion chain also contribute to experimental errors. These error sources constitute control problems that can largely be solved through technical improvements to a few key elements of the apparatus, such as the beam delivery and laser repetition rate.

When testing pulse solutions for parallel gates, as well as for our previously demonstrated two-qubit XX gates, some pulse solutions show inconsistencies between the empirically observed gate performance and the theoretical prediction, with fidelities noticeably worse than expected, even given the experimental error sources, whereas other gate solutions perform as expected; solutions in the latter category are used here. This may be due to non-ideal mode couplings arising from anharmonicities observed in our blade trap, which may be caused by imperfections in the manufacturing and assembly process. It is possible that improvements in trap manufacturing technology, particularly for microfabricated surface traps, may eliminate this issue.

Additional parity curves and fidelity data for two-qubit entangling gates. Additional parity curves and corresponding gate fidelities are shown in Extended Data Fig. 1, with typical fidelities of 96%–99%. An exception is the {(1, 2), (4, 5)} gate, for which the (4, 5) gate has a fidelity of 91% (Extended Data Fig. 1d); however, its phase-space closure diagram in ref. [30] shows that this low fidelity is due to the pulse solution found not being ideal.

Fidelity of parallel two-qubit entangling gates with different degrees of entanglement. Because the XX gates in this parallelization scheme have independent calibrations (see section 'Independence of parallel-gate calibration'), the χ parameters of the two XX gates are independent. The continuously varying parameter χ is directly related to the amount of entanglement generated between the two qubits, given by

$$\mathrm{XX}(\chi)|00\rangle = \frac{1}{\sqrt{2}}[\cos(\chi)|00\rangle - i\sin(\chi)|11\rangle] \quad (4)$$

and can be adjusted in the experiment by scaling the power of the overall gate. Consequently, we can simultaneously implement two XX gates with different degrees of entanglement, which may prove useful for some applications. For example, the full-adder implementation described in the main text requires simultaneously applying an XX($\pi/4$) gate on one pair of qubits and an XX($\pi/8$) gate on another pair of qubits. To demonstrate this capability, Extended Data Fig. 2 shows parity scan data for a simultaneous XX($\pi/4$) gate on ions (1, 5) and an XX($\pi/8$) gate on ions (2, 4). The data are analysed as in Fig. 2 and Extended Data Fig. 1—but we use equation (29) in Supplementary Information (setting $\chi = \pi/4$) to calculate the fidelity for the (1, 5) gate, and equation (28) in Supplementary Information and $\chi = \pi/8$ for the (2, 4) gate. The respective gate fidelities are therefore 96.4(3)% and 99.4(3)%, with an average crosstalk error of 2.2(3)%.

Independence of parallel-gate calibration. Parallel gates can be calibrated independently from one another by adjusting a scaling factor that controls the overall power on the gate without modifying the pulse shape. Furthermore, adjusting a scaling factor that controls the power on a single ion only affects the gate in which the ion participates by modifying the total amount of entanglement, without any apparent ill effects on the gate quality. This is confirmed experimentally using parallel operations on ions (1, 2) and (3, 4) by scanning over the scaling factors associated with ions 1 and 2. Extended Data Fig. 3a, b shows two such scans over the scaling factors for ions 1 and 2 while keeping the (3, 4) gate 'on', with the scaling factor for those two ions set near a fully entangling gate. Extended Data Fig. 3a shows a scan of the scaling factor for only ion 1 while holding the scaling factor for ion 2 constant, and Extended Data Fig. 3b shows a scan over the scaling factor for ions 1 and 2 together. Extended Data Fig. 3c, d shows scans over the scaling factors for ions 1 and 2 while keeping the interaction on (3, 4) 'off'; the scaling factor for the (3, 4) gate is set to 0, so the ions see no light and therefore do not interact during the gate. Extended Data Fig. 3c scans the scaling factor for only ion 2 while holding the scaling factor for ion 1 constant, and Extended Data Fig. 3d shows a scan of the overall scaling factor for ions 1 and 2 together. For all of these scans, as the scaling factors are increased, the population in $|11\rangle$ for ions 1 and 2 increases (and the population in $|00\rangle$ decreases correspondingly), whereas the $|00\rangle$ and $|11\rangle$ populations for the (3, 4) gate remain unchanged.

Optical-power requirements. Although the gate time $\tau_{\mathrm{gate}} = 250\ \mu\mathrm{s}$ for running two XX gates in parallel is comparable to that of a single XX gate (and consequently, comparable to half of the time required to execute two XX gates in series), the parallel-gate scheme requires more optical power. Here, we compare the optical power required for parallel and sequential gates while holding the time per operation constant. The Rabi frequency Ω is proportional to the square root of the beam intensity I, $\Omega \propto \sqrt{I_0 I_1}$, where I_0 and I_1 are the beam intensities for the individual and global beams, respectively. We can therefore calculate the ratio $R_{||}$ of the power required for a gate operation executed in parallel to the power required for a single XX gate on the same ions as $R_{||} = \frac{P_{||}}{P_{\mathrm{XX}}} = \frac{I_{||}}{I_{\mathrm{XX}}} = \left(\frac{\Omega_{||}}{\Omega_{\mathrm{XX}}}\right)^2$. Intensity is power per unit area and, because the beam sizes do not vary, the areas cancel out. The measured power ratios for each experimentally implemented gate are shown in Extended Data Table 1. The power measured is the total optical power that must be generated to apply the gates, regardless of how efficiently that power is used.

Whereas some parallel gates require substantially more power (for example, we had trouble finding a high-quality and low-power solution for {(1, 2), (3, 4)}), most gate operations performed in parallel require about two to four times more power than their single counterparts. We note that the (1, 3) half of the {(1, 3), (2, 5)} parallel gate requires slightly less power than its sequential counterpart; this is probably coincidental, as power minimization is taken into account differently when solving for the sequential two-qubit gate solutions than it is for the parallel-gate solutions. However, a full accounting of the power requirements in this experiment must also take into account power wasted by unused beams and the total time required to perform equivalent operations. Because the individual addressing system has all individual beams on at all times, and these are dumped after the acousto-optic modulator when not in use (see refs [16,30]), any ion that is not illuminated corresponds to an individual beam wasting power. Running two XX gates in parallel takes $\tau_{\mathrm{gate}} = 250\ \mu\mathrm{s}$ and uses beams, each with power P, to illuminate four ions, but performing the same two gate operations in series using stand-alone XX gates requires time $2\tau_{\mathrm{gate}}$ and uses four beams, each with power $P/4$ to $P/2$, to illuminate two ions, wasting two beams. Keeping the time per operation constant, this yields a tradeoff between using twice (or more) the power in half the time versus half the power in twice the time; these parallel gates are then very useful when one has more laser power than time.

Optimized adder circuit. The optimized full-adder circuit implemented in the experiment, shown in Extended Data Fig. 4, is constructed from the circuit in Fig. 4b by combining the CNOT, $C(V)$ and $C(V^\dagger)$ gates from figure 5.12 of ref. [30] and further optimizing the rotations as per the method described in section 5.2.1 of ref. [30]. The two parallel two-qubit operations are outlined in dashed boxes.

The $C(V)$ and $C(V^\dagger)$ gates are the square root of the CNOT gate and its complex conjugate, where $C(V)^2 = C(V^\dagger)^2 = \mathrm{CNOT}$. Consequently, these operations require a two-qubit gate that is the square root of the XX($\pi/4$) gate used for the CNOT gate, which can be achieved with a partially entangling XX($\pi/8$) gate. The unitary for the $C(V) = \sqrt{\mathrm{CNOT}}$ gate is

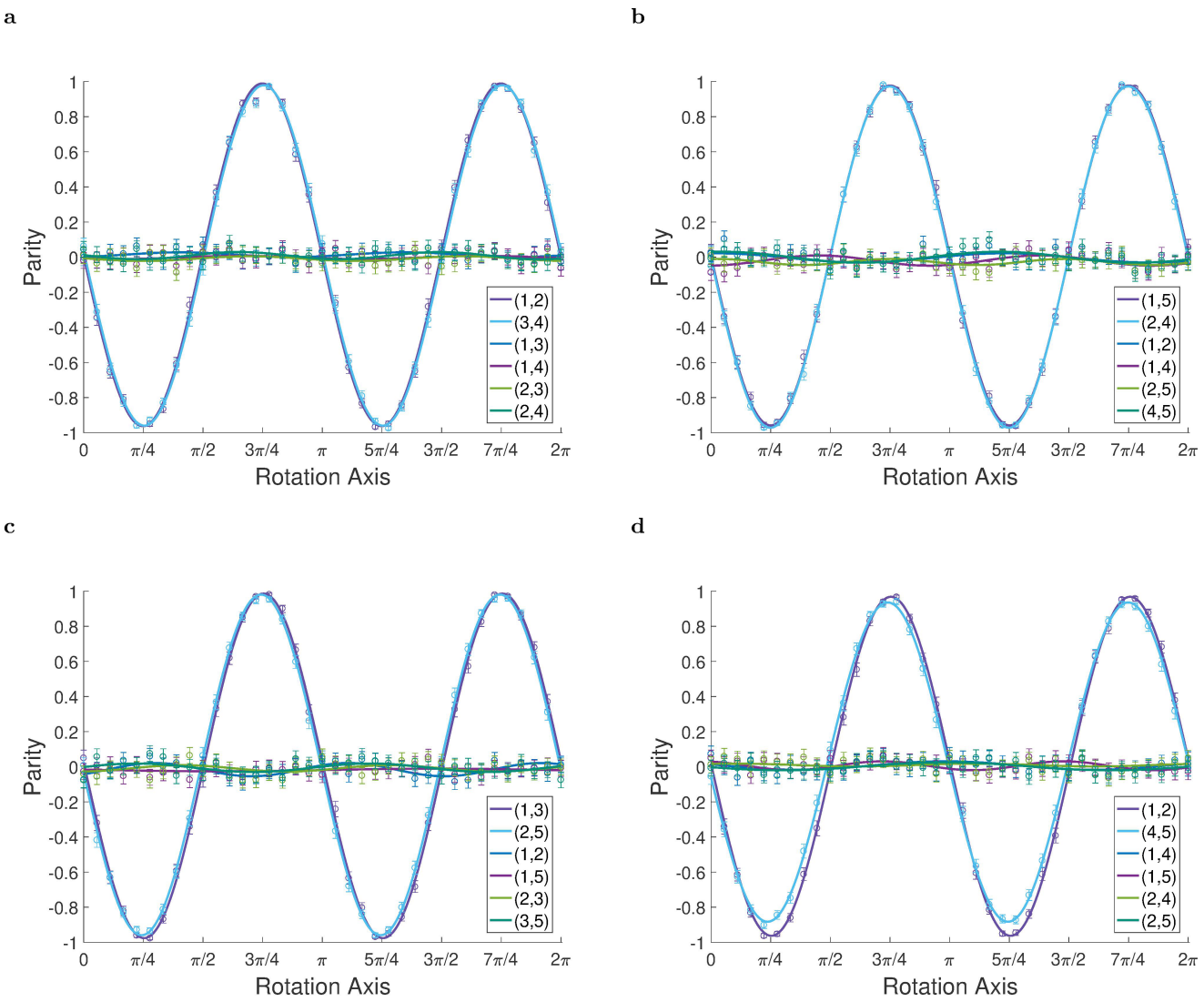
$$U_V = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2}(1-i) & \frac{1}{2}(1+i) \\ 0 & 0 & \frac{1}{2}(1+i) & \frac{1}{2}(1-i) \end{pmatrix} \quad (5)$$

An implementation using XX and R gates is shown in Extended Data Fig. 5. Additional details are available in section 5.9 of ref. [30].

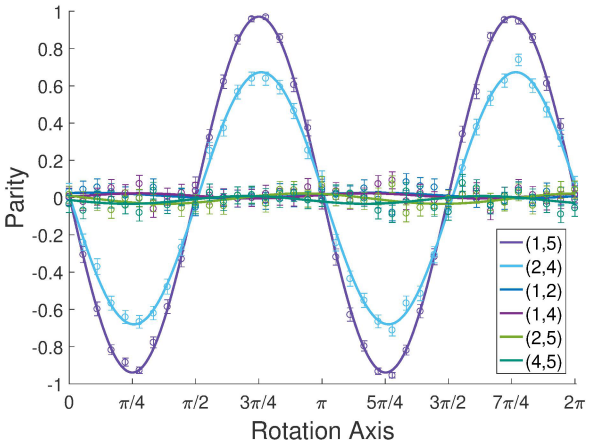
Data availability

All relevant data are available from the corresponding author upon request.

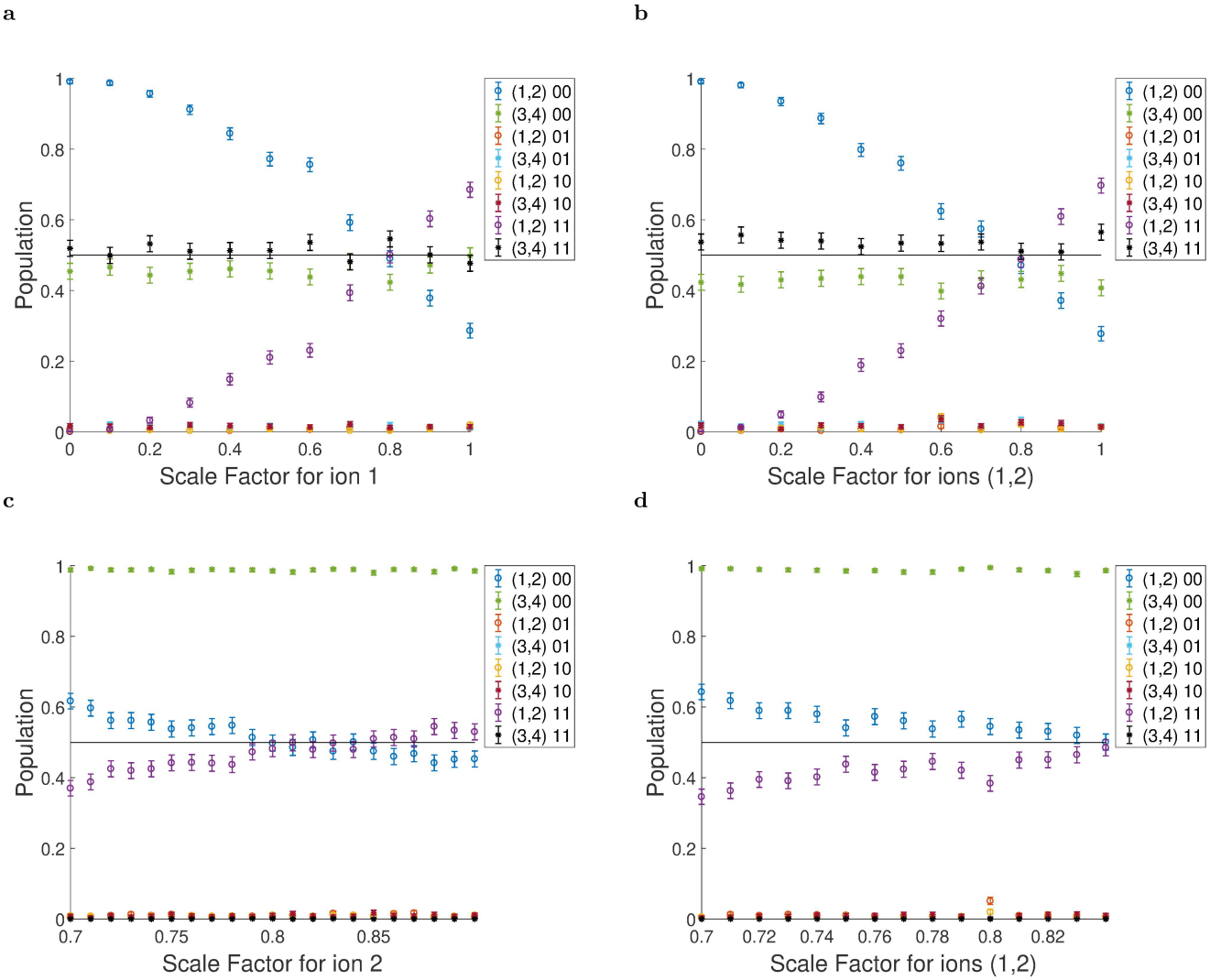
34. Olmschenk, S. et al. Manipulation and detection of a trapped Yb^+ hyperfine qubit. *Phys. Rev. A* **76**, 052314 (2007).



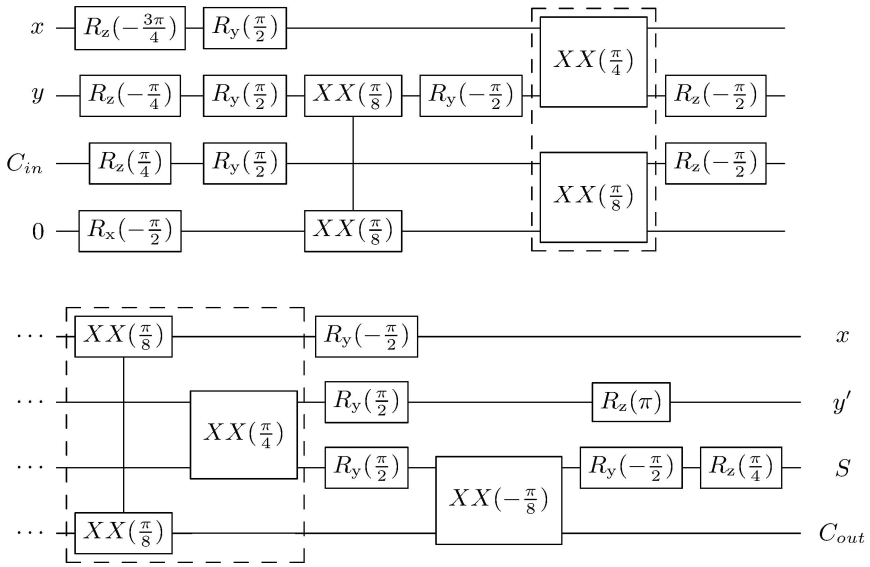
Extended Data Fig. 1 | Additional experimental gate fidelities for parallel two-qubit entangling gates. a–d, Parity curves used to calculate fidelities for parallel XX gates applied on several sets of ions. Circles indicate data, with matching-colour lines indicating calculated fits. The key specifies the ion pair corresponding to each parity curve. The six parity curves shown in each plot include the two gate ion pairs (the first two ion pairs in the key) and the four crosstalk ion pairs. **a**, Ions (1, 2) and (3, 4) yield fidelities of 98.4(3)% and 97.7(3)% for the respective entangled pairs, with an average crosstalk error of 0.6(3)%. **b**, Ions (1, 5) and (2, 4) yield fidelities of 96.8(3)% and 98.1(2)% for the corresponding entangled pairs, with an average crosstalk error of 1.7(3)%. **c**, Ions (1, 3) and (2, 5) yield fidelities of 98.3(3)% and 97.5(2)% for the respective entangled pairs, with an average crosstalk error of 0.8(4)%. **d**, Ions (1, 2) and (4, 5) yield fidelities of 97.2(3)% and 91.9(3)% for the corresponding entangled pairs, with an average crosstalk error of 0.9(3)%.



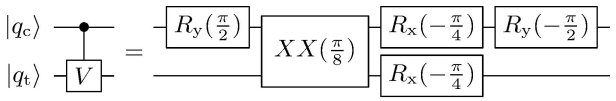
Extended Data Fig. 2 | Experimental gate fidelities for parallel two-qubit partially entangling gates. Parity curve for parallel XX(χ) gates on ions (1, 5) and (2, 4), where an XX(π/4) gate is applied on ions (1, 5) and an XX(π/8) gate on ions (2, 4). Circles indicate data, with matching-colour lines indicating calculated fits. The key specifies the ion pair corresponding to each parity curve. The six parity curves shown include the two gate ion pairs (the first two ion pairs in the key) and the four crosstalk ion pairs. The data yield fidelities of 96.4(3)% and 99.4(3)% for the respective entangled pairs, with an average crosstalk error of 2.2(3)%.



Extended Data Fig. 3 | Independence of parallel-gate calibration. Parallel gates can be calibrated independently. **a**–**d**, Data obtained by applying a pair of entangling gates in parallel and observing the change in population for each pair as the scaling factor for one of the ions or gates is varied. The key specifies the ion pair state corresponding to each dataset; for example, '(1, 2) 00' indicates the 00 population for ions (1, 2). The 01 and 10 populations are very close to 0 and hence not always visible. The error bars are statistical. **a**, Scan of the scaling factor on ion 1 with an entangling gate on ions (3, 4). **b**, Scan of the scaling factor on ions (1, 2) with an entangling gate on ions (3, 4). **c**, Scan of the scaling factor on ion 2 with no light on ions (3, 4). **d**, Scan of the scaling factor on ions (1, 2) with no light on ions (3, 4).



Extended Data Fig. 4 | Full-adder implementation. Application-optimized full-adder implementation using XX(χ), $R_x(\theta)$ and $R_y(\theta)$ gates, where θ is the rotation angle applied by the single-qubit R gate. The two parallel two-qubit operations are outlined in dashed boxes.



Extended Data Fig. 5 | *C*(*V*) gate implementation. Implementation of the $C(V) = \sqrt{\text{CNOT}}$ gate using XX(χ), R$_x$(θ) and R$_y$(θ) gates. The gate is used to construct the full adder used in this work.

Extended Data Table 1 | Comparison of optical power for parallel and single XX gates

Parallel Gate Pairs	$R_{\|\|}$, Pair 1	$R_{\|\|}$, Pair 2
(1,4) and (2,5)	4.3	1.8
(1,2) and (3,4)	7.9	5.0
(1,5) and (2,4)	2.1	1.6
(1,4) and (2,3)	4.3	3.8
(1,3) and (2,5)	0.9	1.5
(1,2) and (4,5)	2.2	2.2

For each pair of parallel XX gates implemented, we compare the optical power required to perform each component XX with its corresponding stand-alone two-qubit XX gate by calculating the power ratio $R_{||}$.

LETTER

https://doi.org/10.1038/s41586-019-1364-3

Committed emissions from existing energy infrastructure jeopardize 1.5 °C climate target

Dan Tong[1,2], Qiang Zhang[2]*, Yixuan Zheng[2,3], Ken Caldeira[3], Christine Shearer[4], Chaopeng Hong[1], Yue Qin[1] & Steven J. Davis[1,2,5]*

Net anthropogenic emissions of carbon dioxide (CO_2) must approach zero by mid-century (2050) in order to stabilize the global mean temperature at the level targeted by international efforts[1–5]. Yet continued expansion of fossil-fuel-burning energy infrastructure implies already 'committed' future CO_2 emissions[6–13]. Here we use detailed datasets of existing fossil-fuel energy infrastructure in 2018 to estimate regional and sectoral patterns of committed CO_2 emissions, the sensitivity of such emissions to assumed operating lifetimes and schedules, and the economic value of the associated infrastructure. We estimate that, if operated as historically, existing infrastructure will cumulatively emit about 658 gigatonnes of CO_2 (with a range of 226 to 1,479 gigatonnes CO_2, depending on the lifetimes and utilization rates assumed). More than half of these emissions are predicted to come from the electricity sector; infrastructure in China, the USA and the 28 member states of the European Union represents approximately 41 per cent, 9 per cent and 7 per cent of the total, respectively. If built, proposed power plants (planned, permitted or under construction) would emit roughly an extra 188 (range 37–427) gigatonnes CO_2. Committed emissions from existing and proposed energy infrastructure (about 846 gigatonnes CO_2) thus represent more than the entire carbon budget that remains if mean warming is to be limited to 1.5 degrees Celsius (°C) with a probability of 66 to 50 per cent (420–580 gigatonnes CO_2)[5], and perhaps two-thirds of the remaining carbon budget if mean warming is to be limited to less than 2 °C (1,170–1,500 gigatonnes CO_2)[5]. The remaining carbon budget estimates are varied and nuanced[14,15], and depend on the climate target and the availability of large-scale negative emissions[16]. Nevertheless, our estimates suggest that little or no new CO_2-emitting infrastructure can be commissioned, and that existing infrastructure may need to be retired early (or be retrofitted with carbon capture and storage technology) in order to meet the Paris Agreement climate goals[17]. Given the asset value per tonne of committed emissions, we suggest that the most cost-effective premature infrastructure retirements will be in the electricity and industry sectors, if non-emitting alternatives are available and affordable[4,18].

International efforts to limit the increase in global mean temperature to well below 2 °C, and to 'pursue efforts' to avoid a 1.5 °C increase, entail a transition to energy systems with net-zero emissions by mid-century[1–5]. Yet recent decades have witnessed an unprecedented expansion of historically long-lived, fossil-fuel-based energy infrastructure—particularly associated with the rapid economic development and industrialization of emerging markets such as China and India[9,10]—and a shift towards natural-gas-fired power plants in the USA. Although existing infrastructure are shown with darker shading, and emissions from proposed power plants (that is, electricity) are more lightly shaded. Numbers within graphs show total amounts of emissions over the period shown.

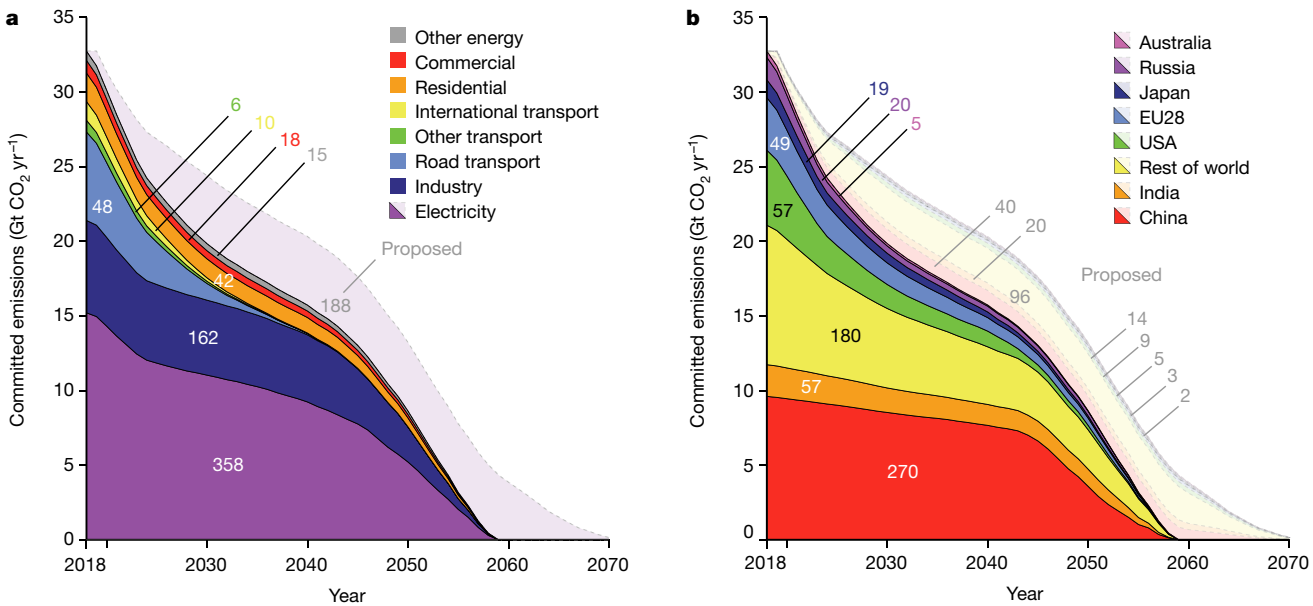


Fig. 1 | Committed annual CO_2 emissions from existing and proposed energy infrastructure. a, **b**, Estimates of future CO_2 emissions by industry sector (**a**; see also Supplementary Tables 1, 2) and country/region (**b**), assuming historical lifetimes and utilization rates. Emissions from

[1]Department of Earth System Science, University of California, Irvine, CA, USA. [2]Ministry of Education, Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing, China. [3]Department of Global Ecology, Carnegie Institution for Science, Stanford, CA, USA. [4]Global Energy Monitor, San Francisco, CA, USA. [5]Department of Civil and Environmental Engineering, University of California, Irvine, CA, USA. *e-mail: qiangzhang@tsinghua.edu.cn; sjdavis@uci.edu

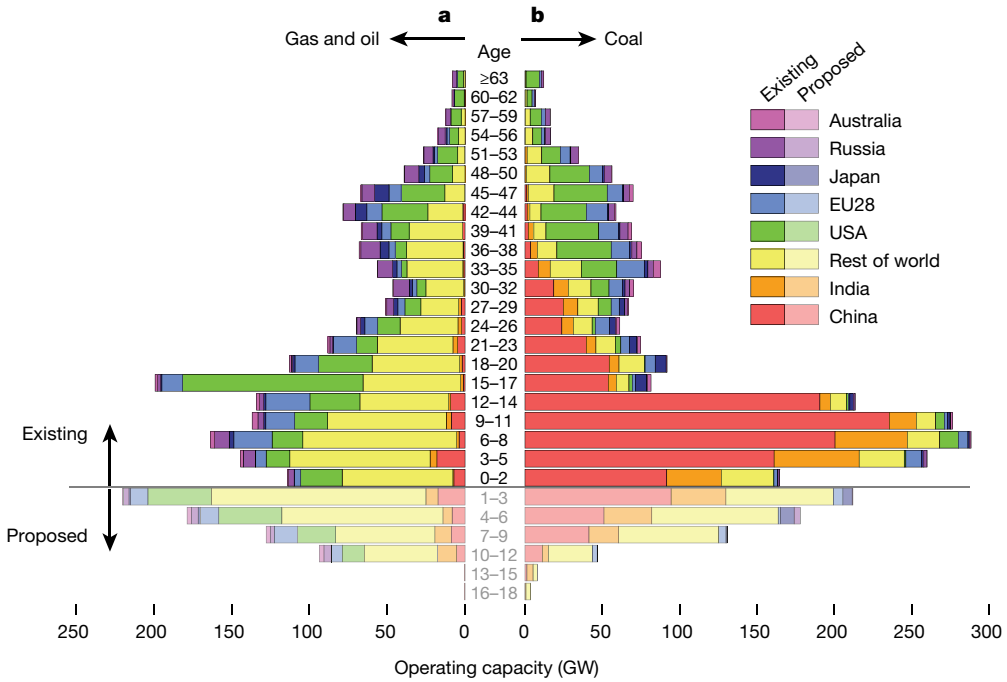


Fig. 2 | Age structure of global electricity-generating capacity. a, **b**, The operating capacity of gas- and oil-fired electricity-generating power units (**a**) and coal-fired units (**b**). The youngest existing units are shown at the bottom of the 'existing' section. The more lightly shaded bars underneath show proposed electricity-generating units according to the year (from now) that they are expected to be commissioned. The recent trends in Chinese and Indian coal-fired units (red and orange at the lower right) and US gas-fired units (green at the left) are easily apparent. '0 years old' means that the power units began operating in 2018.

such expansion may be slowing[19,20], substantial new electricity-generating capacity is proposed—and in many cases is already under construction[12]. Consequently, there is a tension between dwindling carbon-emissions budgets and future CO_2 emissions that are locked-in or 'committed' by existing and proposed energy infrastructure[6,21,22].

A 2010 study estimated that operating fossil-fuel energy infrastructure would emit roughly 500 Gt CO_2 over its lifetime[8]. Subsequent studies estimated that existing power plants alone committed around 300 Gt CO_2 as of 2012 (ref. [9]) and 2016 (ref. [12]), and that existing and proposed coal-fired power plants represented 340 Gt CO_2 as of 2016 (ref. [11]; Extended Data Table 1). Other studies have used integrated assessment models (IAMs) to assess the economic costs of 'unlocking' emissions under stringent climate goals[23,24], and to identify 'points of no return' past which no new infrastructure can be built without exceeding the 2 °C target[25]. Most recently, the potential climate responses to committed emissions were explored[13], using a reduced-complexity climate model and an idealized phase-out of fossil infrastructure to argue that aggressive mitigation of non-CO_2 forcing could yet limit global warming to 1.5 °C. However, it has been nearly a decade since a comprehensive bottom-up assessment of fossil infrastructure and committed emissions was made, during which years China's economy has grown tremendously, there has been a global financial crisis and a natural gas boom in the USA, and the Paris Agreement was ratified and entered into force. Substantial new fossil-fuel energy infrastructure has been commissioned over this period, proposals of new power plants have waxed and waned, and climate-mitigation efforts have grown more ambitious in many countries.

Here we present region- and sector-specific estimates of future CO_2 emissions related to fossil-fuel-burning infrastructure existing and power plants proposed as of the end of 2018, as well as the sensitivity of such estimates to assumed lifetime and utilization rates, and the economic value of associated energy assets. Our analyses are based upon a compilation of the most detailed and up-to-date datasets for energy infrastructure available (see Methods). Our central estimates assume historical lifetimes (for example, 40 years for power plants and industrial boilers and 15 years for light-duty vehicles) and utilization rates (for example, region- and fuel-specific power-plant capacity factors and region-specific averages of vehicle fuel economy and annual kilometres travelled).

Figure 1 shows future CO_2 emissions from existing and proposed energy and transportation infrastructure by sector (Fig. 1a) and country/region (Fig. 1b). We estimate that cumulative emissions by existing infrastructure, if operated as historically, will be 658 Gt CO_2. Of this total commitment, 54% or 358 Gt CO_2 is anticipated to come from existing electricity infrastructure (mainly power plants), reflecting the large share of annual emissions from electricity infrastructure (46% in 2018) and the long historical lifetimes of the infrastructure. Another 25% of the total, or 162 Gt CO_2, is related to industrial infrastructure, and 10% or 64 Gt CO_2 is related to the transportation sector (mainly on-road vehicles; Fig. 1a). This difference reveals the effect of infrastructure lifetimes: although industry and road-transportation sectors have similar annual CO_2 emissions (6.2 Gt and 5.9 Gt CO_2, respectively, in 2018), vehicle lifetimes are roughly a third as long as that of industrial capital. Finally, existing residential and commercial infrastructure represents respectively 42 Gt and 18 Gt CO_2 of committed emissions.

Global committed emissions are now at the apex of a 20-year trend. From 2002 to 2014, as China emerged as a global economic power, total committed emissions grew at an average annual rate of 9% per year (Extended Data Fig. 1a). Meanwhile, committed emissions related to infrastructure in the USA and the 28 member states of the European Union (EU28) have been shrinking since 2006 (Extended Data Fig. 1c). Since 2014, the rate of infrastructure expansion in China and India has also fallen, and committed emissions in China declined by 7% between 2014 and 2018, even as committed emissions in the rest of the world have continued to climb (Extended Data Fig. 1a, c). These most recent trends may reflect nascent shifts in China's economic structure[19] and global trade[20], and may be important harbingers of future changes in regional annual CO_2 emissions[9].

Figure 2 shows the age distribution of electricity-generating units worldwide. Overall, the youth of fossil-based generating units worldwide is striking: worldwide, 49% of the capacity now in operation was

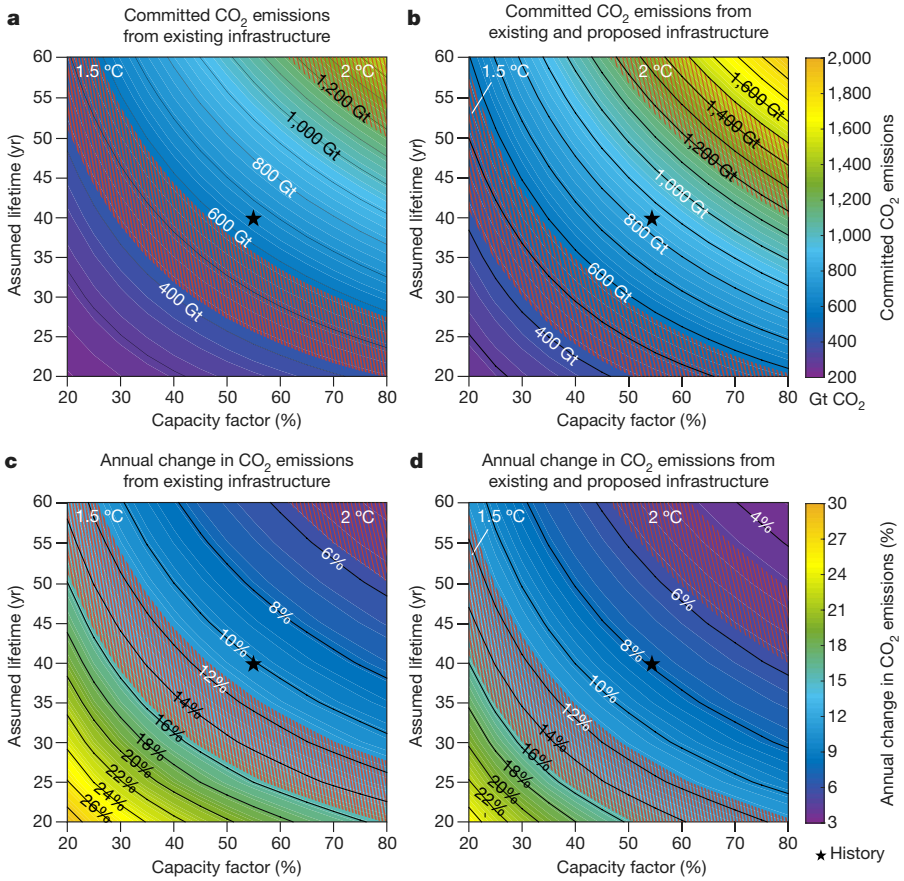


Fig. 3 | Sensitivity of committed emissions and mitigation rates to utilization rates and assumed lifetimes. a, **b**, Committed CO_2 emissions. Contours show estimates of committed CO_2 emissions related to existing infrastructure (**a**) and existing infrastructure plus proposed power plants (**b**) when the assumed lifetimes and utilization rates of electricity and industry infrastructure are varied from 20 years to 60 years (vertical axes) and from 20% to 80% (horizontal axes). **c**, **d**, Committed mitigation rates. For the same ranges of lifetime and utilization as in panels **a**, **b**, the annual rates of emission reduction span from 3% to 30% (**c**, **d**). Hatched orange and red zones indicate carbon budgets and mitigation rates that are likely to limit mean warming to 1.5 °C and 2 °C, respectively (see Methods), and stars denote committed emissions and mitigation rates if existing/and proposed infrastructure is operated as historically.

commissioned after 2004; in China and India, the post-2004 capacity is 79% and 69%, respectively. The average age of coal-fired power plants operating in China and India (11.1 and 12.2 years, respectively) is thus much lower than in the USA and EU28 (39.6 and 32.8 years, respectively; Fig. 2b), with correspondingly longer remaining lifetimes. The predominance of young Chinese infrastructure (which extends to the industrial and transportation sectors; Extended Data Figs. 2, 3) reflects the scale and speed of the country's industrialization and urbanization since the turn of the century. As a result, infrastructural inertia is greatest in China, accounting for 41% of all committed emissions (270 Gt CO_2; Fig. 1b). By comparison, infrastructure in India, the USA and the EU28 represents much smaller commitments: 57 Gt, 57 Gt and 49 Gt CO_2, respectively (Fig. 1b, Supplementary Table 1).

In addition to existing infrastructure, new power plants are being planned, permitted or constructed, and the committed emissions related to such proposed plants can be estimated[11,12]. As of the end of 2018, the best available data showed that 579 gigawatts (GW), 583 GW and 40 GW of coal-, gas- and oil-fired generating capacity respectively was proposed to be built over the next few years (some 20% of it in China; Fig. 2). If built and operated as historically, this proposed capacity would represent an additional 188 Gt CO_2 committed: 97 Gt CO_2 from coal-fired and 91 Gt CO_2 from gas-, oil- and other-fuel-fired generating units (Supplementary Table 2).

Together, committed emissions from existing infrastructure and proposed power plants total 846 Gt CO_2 if all proposed plants are built and all infrastructure is operated as historically (Fig. 1).

Existing electricity and industry infrastructure accounts for 79% of total committed emissions if operated as historically (that is, with a 40-year lifetime and 53% utilization rate; Fig. 1a). However, the lifetime and operation of such infrastructure will ultimately depend on the relative costs of competing technologies, which are in turn influenced by factors such as technological progress and the climate and energy policies in each region[22,26]. Figure 3 highlights the sensitivity of committed emissions (Fig. 3a, b) and the rate of annual emissions reductions (Fig. 3c, d; see Methods) with respect to assumed lifetimes and utilization rates (that is, the capacity factors) of industry and electricity infrastructure (note that the lifetimes and operation of infrastructure in other sectors do not vary from historical averages), with the star in each panel indicating historical average values. For example, total committed emissions related to existing infrastructure decrease to around 200 Gt CO_2 if lifetimes are 20 years and capacity factors are 20%, but increase to almost 1,500 Gt CO_2 if lifetimes and capacity factors are respectively 60 years and 80% (Fig. 3a). These ranges of lifetimes and utilization are quite wide, at the low end probably exceeding economic feasibility for recouping capital investments and covering fixed operating and maintenance costs. When proposed power plants are included, total committed emissions over the same range of lifetimes and capacity factors increase to 263–1,906 Gt CO_2 (Fig. 3b). Maintaining historical capacity factors, a 5-year difference in the lifetime of existing infrastructure represents roughly 70–100 Gt of future CO_2 emissions (Fig. 3a), or about 90–130 Gt if proposed power plants are included (Fig. 3b). Maintaining historical lifetimes and changing the assumed capacity factor by a comparable 9% (for example, from 46% to 55%) results in roughly the same changes in committed emissions, suggesting that these factors have a similar influence.

For comparison, the hatched red and orange zones in Fig. 3a, b show the Intergovernmental Panel on Climate Change (IPCC)'s most recent estimated ranges of remaining cumulative carbon budgets that span the 66%–50% probabilities of limiting global warming to 1.5 °C and 2 °C, relative to the preindustrial era[5]. Excluding proposed power plants, our central estimate of committed emissions (658 Gt CO_2; star in Fig. 3a) exceeds the range of the remaining 1.5 °C budget (420–580 Gt CO_2)[5]. When proposed plants are included, our estimate of committed emissions (846 Gt CO_2; star in Fig. 3b) is two-thirds of the lower estimates of the 2 °C budgets (1,170–1,500 Gt CO_2)[5]. This suggests that, unless compensated by negative-emissions technologies or by retrofitting with carbon capture and storage, 1.5 °C carbon budgets allow for no new emitting infrastructure and require substantial changes to the lifetime

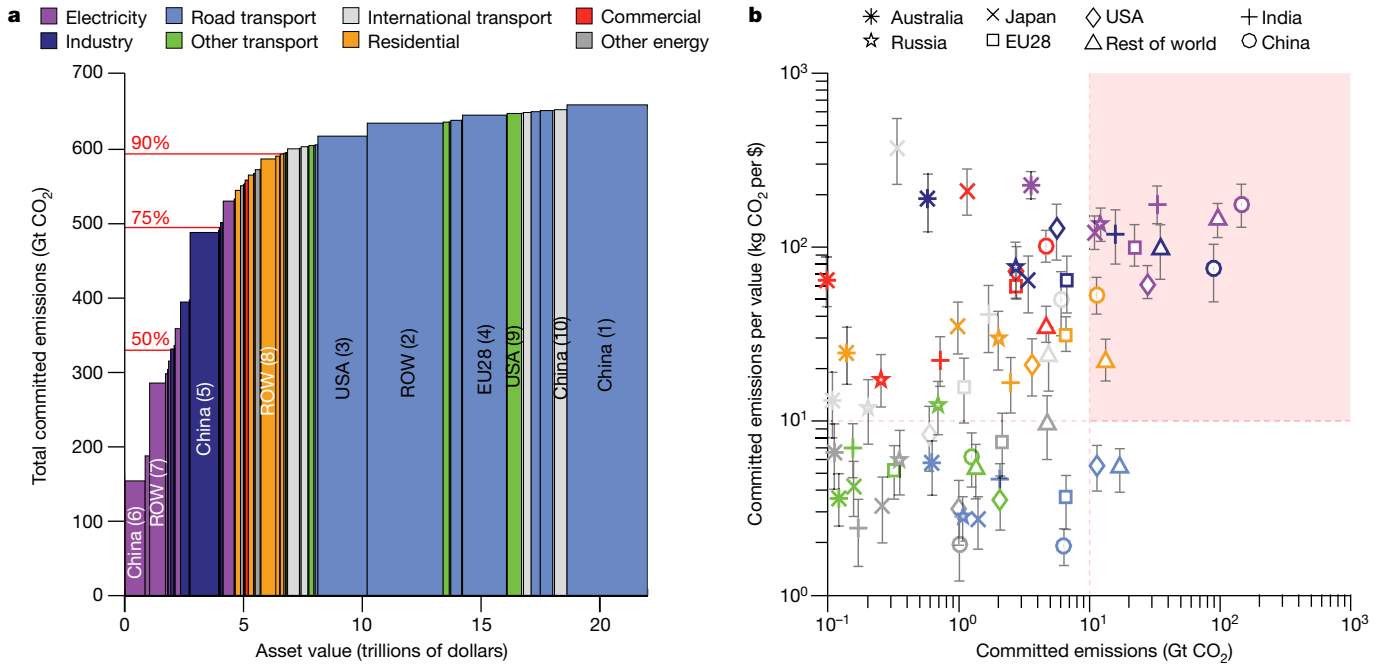


Fig. 4 | Asset value and committed emissions of existing infrastructure. **a**, Rank ordering of CO_2-emitting assets by committed emissions per dollar value reveals large disparities (coloured by sector). The horizontal red lines indicate 50%, 75% and 90% of total committed emissions (658 Gt CO_2) if operated as historically, and the top ten most valuable region sectors are labelled (see Extended Data Fig. 4 for region-specific versions). ROW, rest of world. **b**, Plotting emissions per value (in kilograms of CO_2 per US dollar) against committed emissions suggests targeted opportunities to 'unlock' future CO_2 emissions if alternative technologies become affordable (region sectors in the pink-shaded quadrant). Error bars denote 95% confidence intervals.

or operation of existing energy infrastructure (for example, decreasing lifetimes to less than 25 years or capacity factors to less than 30%; Fig. 3a). Moreover, CO_2 emissions related to the extraction and transport of fossil fuels[27], as well as non-energy CO_2 emissions (for example, resulting from land-use change)[28], are not included in our estimates and will further reduce the remaining carbon budgets.

Climate targets have sometimes been contextualized by the annual rate of emissions reduction they imply. For example, it has been shown[29] that, as of 2013, the cumulative carbon budgets likely to avoid 2 °C of mean warming imply necessary average annual reductions in global CO_2 emissions (that is, mitigation rates) of roughly 6% per year. The hatched areas in Fig. 3c, d show that such mitigation rates, recalculated from the latest carbon budgets, are about 5% per year for the 2 °C budgets (4.5–5.7%) and about 13% per year for the 1.5 °C budgets (11.4–15.7%). By comparison, the contours in the figure show mitigation rates if no new emitting infrastructure is commissioned (10.1%; star in Fig. 3c), or if only already-proposed power plants but no other emitting infrastructure is commissioned (7.9%; star in Fig. 3d). Again, the international targets leave little or no room for new infrastructure if existing plants operate as they have historically (stars), unless fully compensated by negative emissions or retrofitted with carbon capture and storage technology.

Given the constraints of 1.5 °C and 2 °C carbon budgets, we also explore the economic value of existing infrastructure relative to its associated committed emissions. Figure 4a highlights the disproportionality of committed emissions per unit asset value. Together, power and industry infrastructure (purple and dark blue, respectively, in Fig. 4a) represent more than 75% of total committed emissions (519 Gt of 658 Gt CO_2), but less than 25% of the estimated economic value of CO_2-emitting energy infrastructure (roughly US $5 trillion of US $22 trillion; Extended Data Fig. 4 and Supplementary Table 3; see Methods for details of how asset values were amortized). By contrast, transportation infrastructure, with shorter average lifetimes but high capacity costs and a vast number of discrete units, represents roughly two-thirds of the value of emitting assets and less than 10% of committed emissions (Fig. 4a). This analysis suggests that efforts to reduce committed emissions might cost-effectively target the early retirement of electricity and industry infrastructure—despite their often powerful influence on policy and institutions[6,21,22]—if non-emitting alternative technologies are affordable: the magnitude of commitments in these sectors is large and a single dollar of asset value is related to more than 10 kg of future CO_2 emissions (Fig. 4b, red rectangle). Industry and electricity sectors in China represent especially prime targets for unlocking future emissions: nearly half (46%) of these sectors' global committed emissions are associated with Chinese infrastructure (Fig. 4a).

Detailed and up-to-date analysis of existing and proposed CO_2-emitting energy infrastructure worldwide reveals incredibly tight constraints for present international climate targets, even if no new emitting infrastructure is ever built. Although climate and energy analysts have emphasized that avoiding, for example, 1.5 °C of warming remains "technically possible"[5], our results lend vivid context to that possibility: we would have a reasonable chance of achieving the 1.5 °C target with, first, a global prohibition of all new CO_2-emitting devices (including many or most of the already-proposed fossil-fuel-burning power plants); and second, substantial reductions in the historical lifetimes and/or utilization rates of existing industry and electricity infrastructure.

Barring such radical changes, the global climate goals adopted in the Paris Agreement are already in jeopardy and may be contingent upon widespread retrofitting of existing emitting infrastructure with carbon capture and storage technologies (which would be tremendously expensive[30]), large-scale deployment of negative emissions technologies[16], and/or solar-radiation management[4]. On the other hand, our results suggest that the precise level of future warming in excess of the Paris targets depends largely on infrastructure that has not yet been built (Extended Data Fig. 5).

Some important caveats and limitations apply to our findings. The trajectory of future emissions depicted in Fig. 1 represents a scenario in which existing (and proposed) emitting infrastructure 'ages out', and no new emitting infrastructure is ever commissioned. These constraints are not intended to be realistic; rather, they allow

us to isolate and quantify infrastructural—and related economic—lock-in of energy-related emissions[22]. Indeed, technological trends and climate-energy policies that encourage growth in renewable electricity (for example, solar and wind) may lead to early retirement of existing fossil-fuel power plants in some regions (although recent growth of renewable electricity generation has not always displaced fossil-fuel generation[18]). It is also instructive to compare our estimates of committed emissions with plausible energy-emissions scenarios generated by much more sophisticated (but less transparent) IAMs that calculate infrastructure lifetimes and capacity factors endogenously. For example, a recent IAM study of 1.5 °C scenarios found that large-scale CO_2 removal may be necessary to compensate for 'residual' emissions from long-lived and difficult-to-decarbonize sectors of the energy system (for example, freight, aviation and shipping[4])[31].

The size of carbon budgets associated with a given temperature target is also a complicated matter that is sensitive to a host of factors, such as climate sensitivity and non-CO_2 emissions[14,15]. The budgets from the recent IPCC special report[5] are estimates of cumulative net global anthropogenic CO_2 emissions from the start of 2018 until net-zero global CO_2 emissions are achieved (that is, climate is stabilized) with a 66%–50% probability of limiting an increase in mean near-surface air temperatures to 1.5 °C or 2 °C, with limited (less than 0.1 °C) or no overshoot (see Methods for further discussion).

Although ambitious climate targets such as 1.5 °C may help to motivate and accelerate the transition towards net-zero energy systems, their feasibility is often evaluated by the existence of consistent scenarios from IAMs. However, these models have been used to analyse a very large possibility space, and some scenarios may thus reflect aspirational trajectories of energy demand or technological progress and scale whose likelihood may be difficult to evaluate[32,33]. Our data-driven assessment of existing, operating and valuable energy infrastructure may therefore help to elucidate the infrastructural and economic implications of such targets, and also help to identify targeted regional and sectoral opportunities for unlocking future CO_2 emissions.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at https://doi.org/10.1038/s41586-019-1364-3.

Received: 2 December 2018; Accepted: 10 May 2019;
Published online 1 July 2019.

1. Matthews, H. D. & Caldeira, K. Stabilizing climate requires near-zero emissions. *Geophys. Res. Lett.* **35**, L04705 (2008).
2. Meinshausen, M. et al. Greenhouse-gas emission targets for limiting global warming to 2 °C. *Nature* **458**, 1158–1162 (2009).
3. Rogelj, J. et al. Zero emission targets as long-term global goals for climate protection. *Environ. Res. Lett.* **10**, 105007 (2015).
4. Davis, S. J. et al. Net-zero emissions energy systems. *Science* **360**, eaas9793 (2018).
5. Rogelj, J. et al. Mitigation pathways compatible with 1.5°C in the context of sustainable development. In: *Global warming of 1.5 °C. An IPCC Special Report on the Impacts of Global warming of 1.5 °C above Pre-Industrial Levels and Related Global Greenhouse Gas Emission Pathways, in the Context of Strengthening the Global Response to the Threat of Climate Change, Sustainable Development, and Efforts to Eradicate Poverty* (eds Masson-Delmotte, V. et al.) https://www.ipcc.ch/site/assets/uploads/sites/2/2019/05/SR15_Chapter2_Low_Res.pdf (2018).
6. Unruh, G. C. & Carrillo-Hermosilla, J. Globalizing carbon lock-in. *Energy Policy* **34**, 1185–1197 (2006).
7. Matthews, H. D., Gillett, N. P., Stott, P. A. & Zickfeld, K. The proportionality of global warming to cumulative carbon emissions. *Nature* **459**, 829–832 (2009).
8. Davis, S. J., Caldeira, K. & Matthews, H. D. Future CO_2 emissions and climate change from existing energy infrastructure. *Science* **329**, 1330–1333 (2010).
9. Davis, S. J. & Socolow, R. H. Commitment accounting of CO_2 emissions. *Environ. Res. Lett.* **9**, 084018 (2014).
10. Tong, D. et al. Targeted emission reductions from global super-polluting power plant units. *Nat. Sustain.* **1**, 59–68 (2018).
11. Edenhofer, O., Steckel, J. C., Jakob, M. & Bertram, C. Reports of coal's terminal decline may be exaggerated. *Environ. Res. Lett.* **13**, 024019 (2018).
12. Pfeiffer, A., Hepburn, C., Vogt-Schilb, A. & Caldecott, B. Committed emissions from existing and planned power plants and asset stranding required to meet the Paris Agreement. *Environ. Res. Lett.* **13**, 054019 (2018).
13. Smith, C. J. et al. Current fossil fuel infrastructure does not yet commit us to 1.5 °C warming. *Nat. Commun.* **10**, 101 (2019).
14. Rogelj, J. et al. Differences between carbon budget estimates unravelled. *Nat. Clim. Chang.* **6**, 245–252 (2016).
15. Peters, G. P. Beyond carbon budgets. *Nat. Geosci.* **11**, 378–380 (2018).
16. Gasser, T., Guivarch, C., Tachiiri, K., Jones, C. D. & Ciais, P. Negative emissions physically needed to keep global warming below 2 °C. *Nat. Commun.* **6**, 7958 (2015).
17. United Nations Framework Convention on Climate Change *Adoption of the Paris Agreement*. Report No. FCCC/CP/2015/L.9/Rev.1 http://unfccc.int/resource/docs/2015/cop21/eng/l09r01.pdf (UNFCCC, 2015).
18. Le Quéré, C. et al. Drivers of declining CO_2 emissions in 18 developed economies. *Nat. Clim. Chang.* **9**, 213–217 (2019).
19. Guan, D. et al. Structural decline in China's CO_2 emissions through transitions in industry and energy systems. *Nat. Geosci.* **11**, 551–555 (2018).
20. Meng, J. et al. The rise of South–South trade and its effect on global CO_2 emissions. *Nat. Commun.* **9**, 1871 (2018).
21. Erickson, P., Kartha, S., Lazarus, M. & Tempest, K. Assessing carbon lock-in. *Environ. Res. Lett.* **10**, 084023 (2015).
22. Seto, K. C. et al. Carbon lock-in: types, causes, and policy implications. *Annu. Rev. Environ. Resour.* **41**, 425–452 (2016).
23. Bertram, C. et al. Carbon lock-in through capital stock inertia associated with weak near-term climate policies. *Technol. Forecast. Soc. Change* **90**, 62–72 (2015).
24. Johnson, N. et al. Stranded on a low-carbon planet: implications of climate policy for the phase-out of coal-based power plants. *Technol. Forecast. Soc. Change* **90**, 89–102 (2015).
25. Pfeiffer, A., Millar, R., Hepburn, C. & Beinhocker, E. The '2 °C capital stock' for electricity generation: committed cumulative carbon emissions from the electricity generation sector and the transition to a green economy. *Appl. Energy* **179**, 1395–1408 (2016).
26. Wilson, C., Grubler, A., Bauer, N., Krey, V. & Riahi, K. Future capacity growth of energy technologies: are scenarios consistent with historical evidence? *Clim. Change* **118**, 381–395 (2013).
27. Burnham, A. et al. Life-cycle greenhouse gas emissions of shale gas, natural gas, coal, and petroleum. *Environ. Sci. Technol.* **46**, 619–627 (2012).
28. Arneth, A. et al. Historical carbon dioxide emissions caused by land-use changes are possibly larger than assumed. *Nat. Geosci.* **10**, 79–84 (2017).
29. Raupach, M. R. et al. Sharing a quota on cumulative carbon emissions. *Nat. Clim. Chang.* **4**, 873–879 (2014).
30. Rubin, E. S. & Zhai, H. The cost of carbon capture and storage for natural gas combined cycle power plants. *Environ. Sci. Technol.* **46**, 3076–3084 (2012).
31. Luderer, G. et al. Residual fossil CO_2 emissions in 1.5–2 °C pathways. *Nat. Clim. Chang.* **8**, 626–633 (2018).
32. Rogelj, J. et al. Energy system transformations for limiting end-of-century warming to below 1.5 °C. *Nat. Clim. Chang.* **5**, 519–527 (2015); corrigendum **6**, 538 (2016).
33. Grubler, A. et al. A low energy demand scenario for meeting the 1.5 °C target and sustainable development goals without negative emission technologies. *Nat. Energy* **3**, 515–527 (2018).

Acknowledgements D.T. was supported by NASA's Interdisciplinary Research in Earth Science (IDS) programme (80NSSC17K0416) and the National Natural Science Foundation of China (41625020). Q.Z. was supported by the National Natural Science Foundation of China (41625020). C.H., Y.Q. and S.J.D. were supported by the US National Science Foundation (Innovations at the Nexus of Food, Energy and Water Systems (INFEWS) grant EAR 1639318).

Reviewer information *Nature* thanks Gunnar Luderer, Katsumasa Tanaka and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Author contributions S.J.D., D.T. and Q.Z. designed the study. D.T. performed the analyses, with support from Q.Z., Y.Z. and C.S. on datasets, and from S.J.D., Q.Z., K.C., C.H. and Y.Q. on analytical approaches. D.T. and S.J.D. led the writing with input from all coauthors.

Competing interests The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41586-019-1364-3.
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1364-3.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to Q.Z. or S.J.D.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

METHODS

Committed emissions from existing and proposed infrastructure. We extend the approach of ref. [9] to quantify the committed emissions from existing energy infrastructure by integrating more-detailed and up-to-date available data on energy infrastructure, including country- and duty-specific vehicle sales data, and unit-level details on global power plants and Chinese cement kilns and blast furnaces[10,34–39]. We also estimate committed emissions from proposed power plants by collecting information on all proposed power generators from the latest available databases[34,37], in recognition of substantial changes in the pipeline of planned power plants (especially coal) in recent years[34]. Energy infrastructure as quantified in this study is categorized into eight sectors: (1) electricity, (2) industry, (3) road transport, (4) other transport, (5) international transport, (6) residential, (7) commercial and (8) other energy infrastructure (see Supplementary Tables 4 and 5).

Electricity infrastructure. Emissions from electricity infrastructure in this study include all emissions under category 1A1 of the IPCC's revised guidelines[40]. Electricity infrastructure here mainly includes main activity electricity and heat production (1A1a) and petroleum refining (1A1b), as well as the manufacturing of solid fuels and other energy industries (1A1c) (Supplementary Table 5).

Emissions intensities of electricity infrastructure. Previously, we built and published a comprehensive global thermal power plants database (named the Global Power Emissions Database, or GPED) of the year 2010 by integrating high-quality national databases (from China, India and the USA)[10]. Here we update the GPED database to the year 2018 (named GPED-2018) using the latest power plant database from China (CPED)[36] and the Platts World Electric Power Plant (WEPP) database for other regions[37], including all retired and operating units through to the end of 2018. We obtain data and estimates of unit-based CO_2 emission intensity (that is, grams CO_2 per kilowatt-hour) for all units that were operating in 2010 from GPED. For units retired before 2010 or commissioned since 2010, we estimate unit-level CO_2 emission intensity by the methods of ref. [9] on the basis of the Carbon Monitoring for Action (CARMA) database[35] (for older units), or else use national or regional average CO_2 emission intensity for units with the same fuel type and similar nameplate capacity. As prior studies have done, we assume these emissions intensities are constant over a unit's lifetime[8,9].

Assumed lifetime of electricity infrastructure. In the resulting GPED-2018, the global average lifetimes of retired coal-, natural-gas- and oil-fired power units are 35.9, 37.1 and 33.9 years, respectively. Consistent with ref. [9], we simplify these ranges to a single reference lifetime of 40 years for all electricity-generating units for our 'as historically' case, and show the sensitivity of committed emissions to this assumption in Fig. 3. When units are already operating beyond their assumed lifetime, we randomly retire them over the next five years in order to avoid unrealistically abrupt changes in emissions between 2018 and 2019.

In addition, we assume that the age structure and lifetime of autoproducers (industrial and commercial facilities that generate their own electricity on-site)[40] and other energy industries are similar to the main-activity power plants in each region. Therefore, committed emissions from existing electricity infrastructure are quantified by using the survival curves derived from main-activity power plants, scaled to include these other types of electricity infrastructure by using country-level electricity emissions totals in 2018 from the International Energy Agency (IEA). Note that, because of data availability[41], we derived the country-level CO_2 emissions from fossil-fuel combustions for 2018 by multiplying country-level CO_2 emissions in 2016 by projected change rates during 2016–2018.

Finally, we quantify cumulative future CO_2 emissions from proposed power plants by the same procedure (assuming historical average unitization rates and lifetimes), using a database of proposed coal-fired units that has been developed by CoalSwarm[34] and the planned units fired with other fossil fuels from the 2018 (fourth quarter) WEPP database[37].

Industry infrastructure. Industrial emissions in this study include all emissions under category 1A2 of the IPCC's revised guidelines[40]. For all countries but China, we estimate cumulative future emissions from industry infrastructure by using country-level emissions data for the year 2018 obtained from the IEA, assuming that the age distribution and survival curves of each region's industry infrastructure are consistent with its electricity infrastructure. To derive China's industrial survival curves, we use unit-level details of cement kilns and blast furnaces (iron and steel) that are currently operating in China (Extended Data Fig. 2), obtained from China's Ministry of Ecology and Environment (MEE) (our unpublished data, referred to hereafter as the MEE database).

Our detailed data on Chinese infrastructure represent an important improvement over prior estimates of committed emissions, as China alone accounts for roughly 47% of total industrial emissions[41]. In particular, the iron/steel and non-metallic minerals (for example, cement and glass) industries account for about 50% of all industrial CO_2 emissions in recent years[41], and China produced 49.6% of the world's raw steel and 57.3% of the world's cement in 2016 (ref. [42]). The unit-level data on China's industrial infrastructure thus substantially decrease the uncertainty of committed industry emissions, by alleviating the need for assumptions related to almost half of global industry infrastructure (that is, 9.0% of global CO_2 emissions from all sources[41]). Moreover, we observe that the age distributions of electricity and industry infrastructure in China are quite similar (Extended Data Fig. 6), which lends support to our assumption that this is the case in other regions for which we lack detailed data on industrial infrastructure.

Transportation infrastructure. Transport emissions in this study include all emissions under category 1A3 of the IPCC's revised guidelines[40], which includes emissions from road transport, other transport and international transport (Supplementary Tables 4, 5).

We calculated cumulative future emissions from road transport following the approach in ref. [8] and further updating the activity rates with updated country-, region- and duty-specific vehicle sales data[38,39] (that is, 18% of global CO_2 emissions from all sources[41]). Specifically, we use the number, class and vintage of motor vehicles sold during 1977–2017 from 40 major countries and regions[38,39] (information for 2018 was derived by projecting 2016–2017 rates of change one additional year; Extended Data Fig. 3). Owing to data availability, we estimate the number of vehicles remaining on the road over time by using class and model year-specific survival rates of US and Chinese vehicles to represent developed (the USA) and developing (China) countries or regions[43,44]. We then calculate annual vehicle emissions by using the average miles driven per year (MPY) per vehicle by class, and carbon emission factors of 10.23 kg and 11.80 kg CO_2 per gallon of gas and diesel, respectively, and scale our estimated emissions to match country-level road-transport emissions in 2018 as reported by the IEA[41].

'Other transportation' infrastructure includes existing aviation, rail, pipeline, navigation and other non-specified transport. International transport infrastructure includes international marine bunkers and international aviation bunkers (Supplementary Table 4). Again, we follow ref. [8], estimating cumulative future CO_2 emissions from existing other and international transport by using country-level emissions data for 2018 from IEA, and assuming lifetimes and age distributions similar those of to motor vehicle fleets in each country/region.

Residential, commercial and other energy infrastructure. Residential and commercial emissions are included under category 1A4 of the IPCC's revised guidelines[40], and 'other energy' emissions include, for example, emissions from agriculture, forestry, fishing and aquaculture under category 1A4, as well as stationary, mobile and multilateral operations under category 1A5. We calculated cumulative future emissions from this infrastructure by using country-level emissions data for 2018 derived from the IEA[41], and assuming that age distributions and lifetimes of residential, commercial and other energy infrastructure in each region were similar to electricity infrastructure in the same region in the absence of better information.

The least-supported methodological assumptions that we make thus concern this residential, commercial and other energy infrastructure (representing around 10% of total fossil fuel CO_2 emissions in 2016; ref. [41]), where we lack any unit-level data. In order to test the sensitivity of total committed emissions from this infrastructure, we performed additional analyses of different assumed lifetimes. We found the committed emissions from residential, commercial and other energy infrastructure to be 29, 74 and 135 Gt CO_2 when lifetimes of respectively 20, 40 and 60 years are assumed (Extended Data Fig. 7). That is, our estimates of total committed emissions from all existing energy infrastructure decrease by 7% (to 613 Gt CO_2) if lifetimes of residential, commercial, and other energy infrastructure are assumed to be 20 years, and increase by 9% (to 719 Gt CO_2) if the lifetimes are assumed to be 60 years. In comparison with the carbon budgets associated with targets of 1.5 °C and 2 °C, these are relatively small effects, and not substantial enough to affect the main conclusions of our study.

Comparison of cumulative future emissions estimates. Other studies[8,9,11–13] have analysed committed emissions from various infrastructures in different ways, as mentioned in the text and summarized in Extended Data Table 1.

For example, refs [11,12] both reported committed emissions relating to existing and planned power plants using 2016 data. Although the latter analysed committed emissions from all fossil electricity infrastructure[12], the former focused particularly on coal-fired units[11]. Importantly, the 2018 data used herein reveal that substantial cancellations of proposed plants have occurred over the intervening two years: whereas the previous studies estimated that around 150 Gt CO_2 (ref. [11]) and 210 Gt CO_2 (ref. [12]) were committed by proposed coal plants, we estimate only around 100 Gt CO_2—that is, 50–100 Gt CO_2 less (or 10%–20% of the remaining carbon budget that is consistent with 1.5 °C warming). Moreover, our study contains more-detailed estimates of regional commitments and the sensitivity of these commitments to assumed lifetime and capacity factor.

Most recently, ref. [13] estimated the global warming related to committed emissions by using a reduced-complexity climate model (Finite Amplitude Impulse Response, or FaIR). Their study also included estimates of committed emissions from all sectors, but these relied on past estimates of the age distribution of fossil-fuel infrastructure and an idealized, linear phase-out of such infrastructure[13]. Because turnover of infrastructure has decreased the median

age of electricity-generating capacity in many regions (Fig. 2), our estimates of electric power sector commitments (358 Gt CO_2) are about 13 Gt CO_2 greater than those used in ref. [13] (345 Gt CO_2). Our data-driven approach also permits region-specific results, analysis of the trend in commitments over time, inclusion of proposed power plants, and an assessment of the economic value of underlying infrastructures. Yet, because the estimates of CO_2 emissions committed by other infrastructure in ref. [13] are larger than our bottom-up estimates (Extended Data Table 1), the overall estimate reached by their idealized approach (715 Gt CO_2) is nonetheless similar to ours (658 Gt CO_2).

The authors of ref. [13] assess global climate responses to committed CO_2 increases and conclude that the world is not yet committed to a 1.5 °C warming. However, it is difficult to directly compare the magnitude of the CO_2 emissions in the phase-out scenarios of ref. [13] with the 1.5 °C carbon budgets in the IPCC's special report[5] (SR1.5), for two reasons. First, although SR1.5 also used the FaIR model in its procedure for evaluating non-CO_2 forcing, it did not use the FaIR model's transient climate response to cumulative emissions (TCRE), which is smaller and would have led to considerably larger carbon budgets. Second, the mitigation scenarios evaluated ref. [13] also assumed that non-CO_2 emissions are completely phased out in parallel to CO_2 emissions, but the integrated assessment model scenarios on which SR1.5's non-CO_2 forcing (and carbon budgets) are based do not completely eliminate non-CO_2 emissions this century[45].

Variation in utilization rates and assumed lifetimes. As described above, cumulative future committed emissions from electricity and industry infrastructure depend on present utilization rates and assumed lifetimes. The longer the assumed lifetime and higher the utilization, the greater the estimate of committed emissions will be. Therefore, we test the sensitivity of committed emissions to assumed lifetimes and utilization rates of energy and industry infrastructure across lifetimes from 20 years to 60 years, and utilization rates of 20% to 80%.

Remaining carbon budgets to limit mean warming to 1.5 °C and 2 °C. As described in the text and discussed in recent literature, the size of carbon budgets associated with a given temperature target is a complicated matter that is sensitive to a host of factors[14,15], including: (1) whether the budget reflects cumulative net emissions until the temperature target is exceeded, or cumulative net emissions that limit the global temperature increase to below the target (that is, climate is stabilized); (2) whether there can be a temporary overshoot of the temperature target (and by how much)[46]; (3) the climate responses to CO_2 and non-CO_2 forcings[47]; (4) the magnitude and Earth-system response to negative emissions[48]; (5) how global temperature is calculated; (6) the pre-industrial baseline used[49]; (7) whether Earth-system feedbacks such as permafrost thawing are included[50–53]; and (8) future emissions of non-CO_2 greenhouse gases and aerosols[54,55].

The magnitude of non-CO_2 forcing is particularly relevant to assessments of committed emissions, because non-CO_2 forcing is inversely related to the remaining carbon budget[54,55], and because some non-CO_2 greenhouse gases and aerosols are directly related to the current energy system (for example, fugitive methane[56]) or are co-emitted with CO_2 by fossil-fuel-burning infrastructure. Other large sources of non-CO_2 gases and aerosols exist outside of the energy system, such as agriculture[57]. For the SR1.5 budgets[5], non-CO_2 forcing was estimated using integrated assessment model scenarios and a pair of reduced-complexity climate models (Model for the Assessment of Greenhouse-gas Induced Climate Change (MAGICC) and FaIR), with substantial uncertainties associated with both scenario variations (± 250 Gt CO_2) and climate responses (−400 Gt to 200 Gt CO_2) for the 1.5 °C budget. Non-CO_2 greenhouse gases and aerosols decline but do not reach zero in any of the scenarios assessed in the SR1.5 report. By contrast, ref. [13] modelled the complete phase-out of non-CO_2 emissions in parallel with energy-related CO_2 emissions—a formidable scenario that was found to have a high probability (64%) of limiting warming to 1.5 °C.

In this study, we compare our estimates of committed emissions to the SR1.5 budgets[5]. As defined in SR1.5, 'remaining' carbon budgets are the cumulative net global anthropogenic CO_2 emissions from a given start date (1 January 2018) to the year in which such emissions reach net zero that would result, at some probability, in limiting global warming to a given level[5]. By this definition, budgets are not simply cumulative emissions until the time at which mean temperature exceeds a given threshold[14], but rather what have been called 'threshold avoidance' or 'stabilization' budgets. The SR1.5 budgets were derived from the transient climate response to cumulative CO_2 emissions in climate model simulations that have been further adjusted to include additional climate forcing related to non-CO_2 greenhouse gases and aerosols[45]. They do not include Earth-system feedbacks (which SR1.5 suggests could reduce the remaining budgets by 100 Gt CO_2 over this century).

However, as remaining budgets associated with a mean surface warming of 1.5 °C dwindle, uncertainties in transient climate responses to CO_2 emissions[15,47] and the current and future non-CO_2 forcing loom large[53–55]. In order to make our results as useful, transparent and comparable as possible, we report positive, CO_2-only commitments from existing and proposed fossil-fuel-burning infrastructure, and compare these to the remaining (stabilization) carbon budgets reported by SR1.5 to give a 66%–50% probability of limiting warming to 1.5 °C and 2 °C with little (0.1 °C) or no overshoot: that is, 420–580 Gt CO_2 and 1,170–1,500 Gt CO_2, respectively (see table 2.2 in ref. [5]). Thus, if not offset by negative emissions, the total committed emissions that we estimate if existing infrastructure operates as it has historically (that is, 658 Gt CO_2) would make it likely that global temperatures will exceed 1.5 °C unless the remaining carbon budgets in SR1.5 are substantially wrong. For example, the climate response to CO_2 could be less than expected on the basis of the climate model simulations assessed in SR1.5, and/or non-CO_2 forcing in the future could be much less than it is on average in the integrated assessment model scenarios that were assessed by SR1.5. Indeed, ref. [13] analysed a future in which both are true.

Estimates of the annual rate of emission reductions. We estimate annual rates of emissions reduction ('mitigation rates') following ref. [29]:

$$f(t) = f_0(1 + (r + m)t)\exp(-mt)$$

where $f(t)$ is the emissions at time t; f_0 is the emissions at the start of mitigation ($t = 0$); r is an initially linear growth rate; m is the annual rate of emission reductions; and r and m both have units of 'per year'. We calculate the annual rate of emission reductions needed to meet a quota, q, from $t = 0$ onward (with emission time $T = q/f_0$) as:

$$m(q) = \frac{1 + \sqrt{1 + \frac{rq}{f_0}}}{\frac{q}{f_0}} = \frac{1 + \sqrt{1 + \frac{r}{T}}}{T}$$

We use initial emissions, f_0, at 2018 (32.7 Gt) and growth rates, r, averaged over 2013–2018 (0.028%) (obtained from the IEA[41]) to estimate mitigation rates under different cumulative CO_2 emissions, which we assumed to be equivalent to the carbon quota, q.

Estimates of asset value from existing infrastructure. We estimate the asset value by sector and by country/region using the following equation:

$$\mathrm{AV}_{i,s} = \sum_{n=\mathrm{PY}-\mathrm{LT}}^{\mathrm{PY}} \sum_{y} \{\mathrm{TC}_{i,s,n,y} \times \mathrm{CC}_{i,s,n,y} \times [(1-\mathrm{RV}) \times \mathrm{DR}_{i,s,n,y} + \mathrm{RV}]\}$$

where i, s, n and y represent the country/region, sector, years and combustion/production technology, respectively; AV is the asset value; TC is the equivalent total capacity/numbers; CC is the capital costs; RV is the ratio of residual value, with 5% applied for all infrastructure; DR is the depreciation rate; PY is the present year (2018 in this study); and LT is lifetime.

We adopt a sector-dependent method, and apply straight-line and geometric models for different infrastructures, as in Supplementary Table 6. We collected data on capital costs used to estimate asset values from previous literature[12,21,23–25,58,59] and various reports[60–64]. Wherever possible, we use interannual and national average capital costs for different combustion/production technologies and equipment. Where interannual and national averages are not available, we instead use an average for all of the countries in the same region for which capital cost data are available.

Electricity infrastructure. We estimate the total value of fossil-fuel-based electricity-generating assets according to each unit's power-generating capacity (in kilowatts) and age, as well as fuel- and technology-specific capital costs (in dollars per kilowatt).

The assumed lifetime of coal power plants is 40 years. Although plants can operate for considerably longer periods, shutting down a plant after its assumed lifetime will not result in any stranded capital investment, since the initial capital cost will have been fully paid[24]. Thus, our estimates only include the asset value of operating electricity-generating units that are now less than 40 years old. Unit-level details of electricity-generating technologies were obtained from the GPED-2018 database.

In addition, part of the committed CO_2 emissions in electricity infrastructure is from heating plants. We have evaluated the asset value of combined heat and power (CHP) plants along with that of other power plants, but we estimate the asset value of individual heating plants separately, using IEA data on heating output (in terajoules, TJ)[65,66] to estimate the capacity of such heating plants and converting this to an equivalent power capacity (in GW) by assuming that they operate with the average utilization rates of power-generating units in the same region. Supplementary Table 6 summarizes our assumptions in estimating asset values for individual heating plants.

Industrial infrastructure. 'Industrial infrastructure' includes various facilities and systems from different subindustrial sectors (Supplementary Tables 4 and 5). Considering the difficulty of collecting the operating capacity for all of the subindustrial sectors, we estimate the value of industry infrastructure as the combined asset values of cement, iron and steel plants, and industrial boilers. As described above, we estimated the asset values for cement, iron and steel capacity that has

been operating less than 40 years only. We quantified asset values from the cement, iron and steel industries through total capacity and capital investment per unit (Supplementary Table 6).

We estimate total capacities (in tonnes per hour, t h^{-1}) of industrial boilers at country- or region-specific level by fuel type, using total energy consumptions obtained from the IEA[65,66]. We assume the utilization rates of industrial boilers to be the same as the average utilization rates of electricity infrastructure. The related assumptions are shown in Supplementary Table 6.

Transport infrastructure. We quantify the asset values from road transport, other transport and international transport separately. For road-transport infrastructure, we estimate asset value using the number of annual vehicle sales, annual average new car prices, and a depreciation-rate function. The data sources for the number of annual vehicle sales are described above, and we further collect annual average new car prices by vehicle type and country/region[39]. Because depreciation rates tend to be considerably lower in developing countries than in industrialized countries[67], we adopt different depreciation-rate functions for developing and developed countries[67].

For international-transport infrastructure, we estimate the value of international ships and international airplanes. Owing to limited data availability, we use the same approach as with heating infrastructure, basing our estimates on the total energy consumption (fuels) for international aviation and international navigation from the IEA, and converting to the number of reference narrow-body aircraft and standardized international freight ships by such fuel consumption. Specifically, we assume 2 million kilometres per year for each aircraft, and 149 megajoules per airplane kilometre, for reference narrow-body aircrafts[21] (Supplementary Table 6); and 940 million annual tonnes per kilometre, and an average ship energy intensity of 0.125 megjoules per tonne kilometre, for international freight ships[21]. We use the same total average depreciation rates for international transport as we do for road-transport infrastructure.

We use a similar approach for other transport (that is, domestic ships, domestic airplanes and non-specific transport), adopting the same assumptions applied for international transport for domestic ships and domestic airplanes. For non-specific transport, we quantify asset values by converting to the number of conventional diesel heavy-duty freight trucks. The corresponding assumptions are shown in Supplementary Table 6.

Residential, commercial and other energy infrastructure. We quantify the asset values of residential, commercial and other energy infrastructure separately using sector- and fuel-specific energy-consumption data from the IEA[65,66].

Residential and commercial infrastructure uses energy for space heating, heating water, and cooking. Other energy infrastructure includes uses of energy for agriculture, fishing and other activities. Given very limited data, we quantify the value of residential and commercial infrastructure by using an equivalent capacity of normalized space heating units, water-heating units and cooking equipment. For the 'other energy' infrastructure, we quantify the asset value by converting to normalized agriculture machines, fishing boats and boilers. We then apply the total average depreciation rates of electricity infrastructure to these residential, commercial and other energy infrastructures.

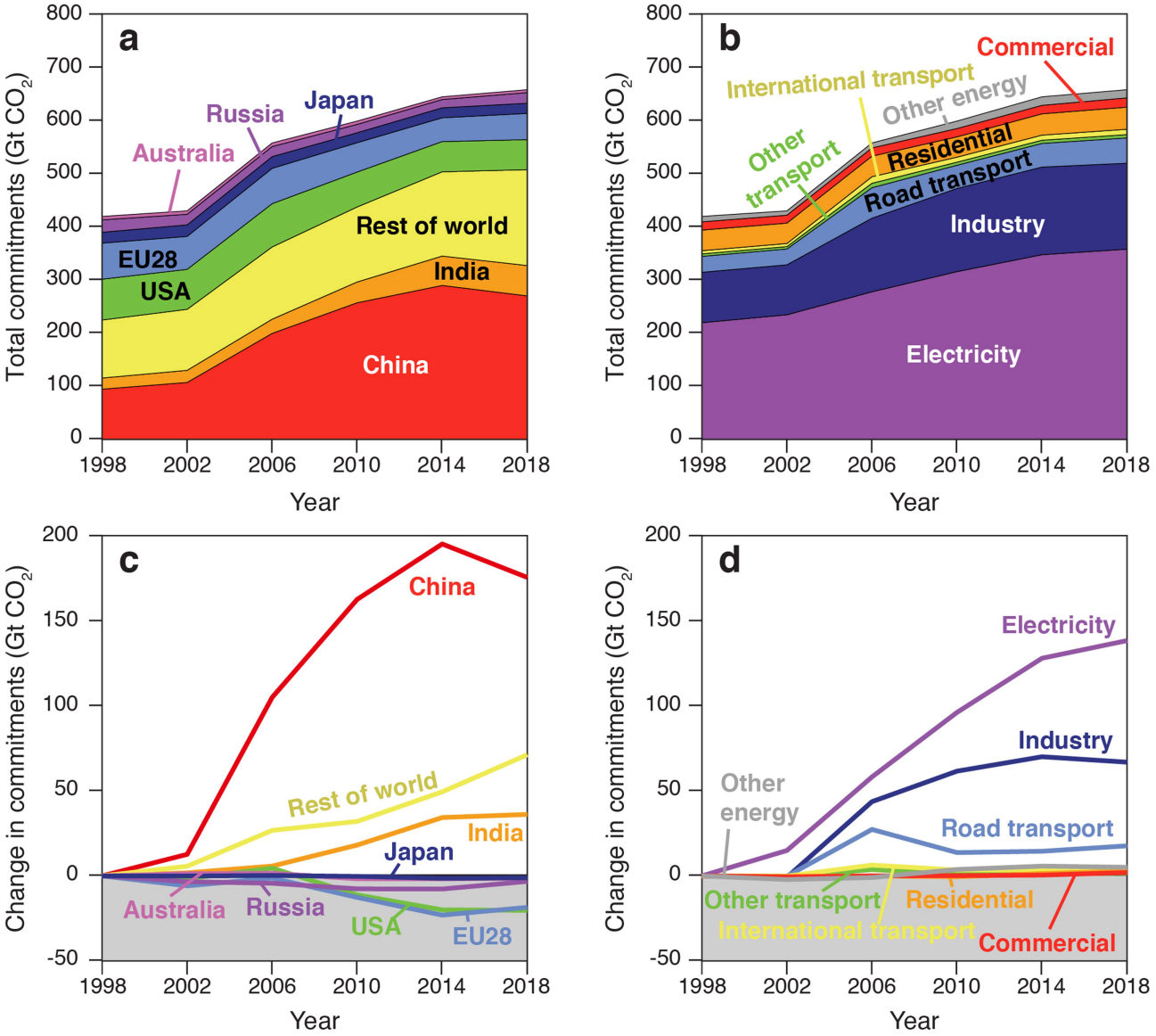
Uncertainty estimation. Our estimates of asset values are subject to uncertainty owing to incomplete knowledge of operating capacities, age structure and capital costs per unit. In order to more completely assess uncertainties in our results, we perform a Monte Carlo analysis of asset values by sector and by country/region, in which we vary key parameters according to published ranges[58,68,69] and collected capital costs data as above. The error bars in Fig. 4 depict the results of this analysis, showing the lower and upper bounds of a 95% confidence interval (CI) around our central estimate. The Monte Carlo simulation uses specified probability distributions for each input parameter (for example, capital cost per unit, and the ratio of residual value) to generate random variables[68]. The probability distribution of asset values is estimated according to a set of runs ($n = 10{,}000$) in a Monte Carlo framework with probability distributions of the input parameters. The ranges of sector and region parameter values vary in part because of the quality of their statistical infrastructures[69]. Supplementary Table 7 summarizes the probability distributions of the asset value estimation-related parameters.

Data availability

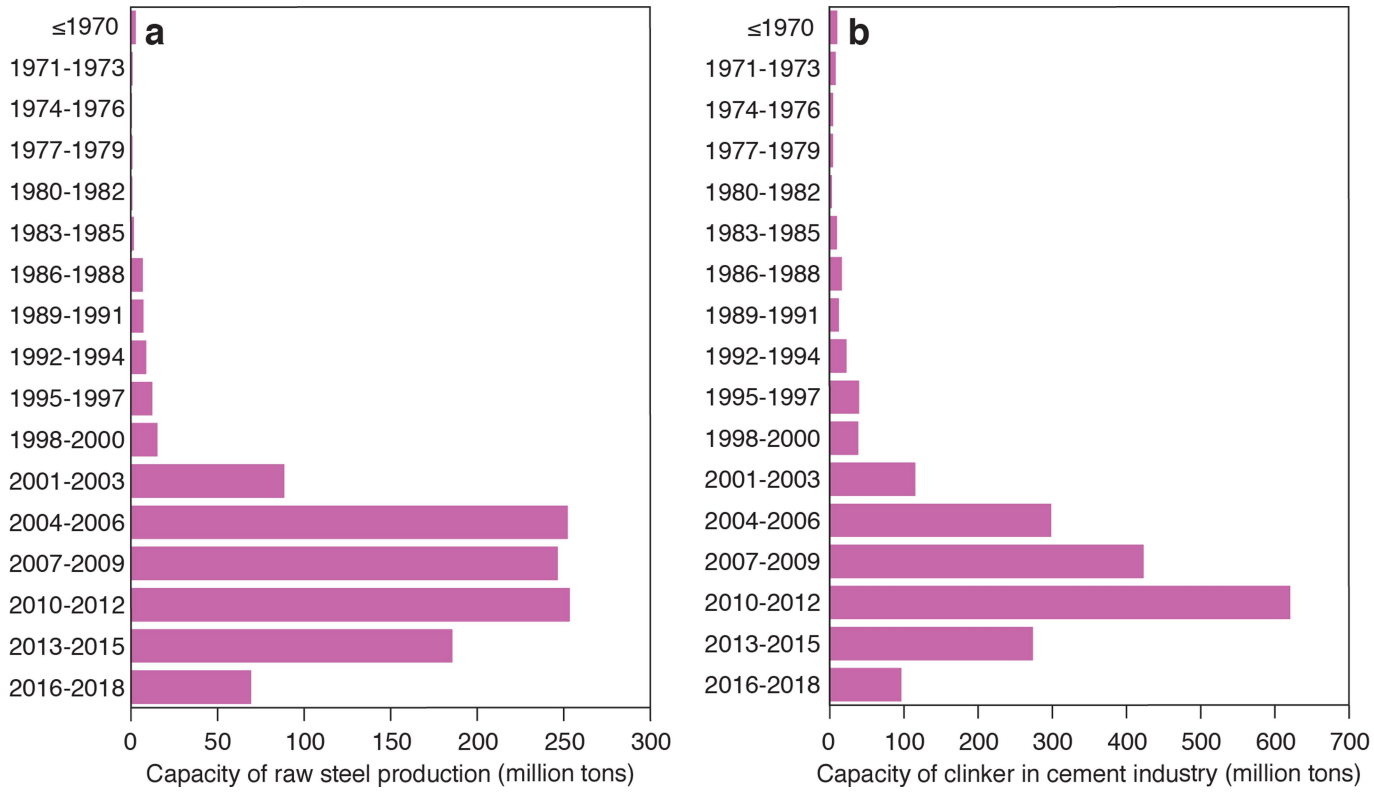
The numerical results plotted in Figs. 1–4 are provided with this paper. Our analysis relies on six different data sets, each used with permission and/or by license. Five are available from their original creators: (1) the GPED database: http://www.meicmodel.org/dataset-gped.html; (2) Platt's WEPP database: https://www.spglobal.com/platts/en/products-services/electric-power/world-electric-power-plants-database; (3) the Carbon Monitoring for Action (CARMA) database: http://carma.org/; (4) the CoalSwarm database: https://endcoal.org/tracker/; and (5) vehicle sales data: https://www.statista.com/markets/419/topic/487/vehicles-road-traffic/. The sixth data set includes unit-level data for Chinese iron, steel and cement infrastructure, which we obtained directly from the Chinese Ministry of Ecology and Environment. We do not have permission to share the raw data, but we provide it in an aggregated form (Extended Data Fig. 2).

34. Shearer, C. et al. *Boom and Bust 2018: Tracking the Global Coal Plant Pipeline*. March 2018 Report (CoalSwarm, Sierra Club and Greenpeace, 2018).
35. Ummel, K. *CARMA Revisited: An Updated Database of Carbon Dioxide Emissions From Power Plants Worldwide*. Working Paper 304 (Center for Global Development, 2012).
36. Tong, D. et al. Current emissions and future mitigation pathways of coal-fired power plants in China from 2010 to 2030. *Environ. Sci. Technol.* **52**, 12905–12914 (2018).
37. World Electric Power Plant Database (WEPP). *S&P Global Platts* https://www.spglobal.com/platts/en/products-services/electric-power/world-electric-power-plants-database (2018).
38. Ward's World Motor Vehicle Data. *Wards Intelligence* http://wardsauto.com/wards-world-motor-vehicle-data-0 (2008).
39. Statistics and Facts about Vehicles & Traffic, 2006-2017. *Statista* https://www.statista.com/ (2018).
40. *IPCC Guidelines for National Greenhouse Gas Inventories* Vol. 4 (Bracknell, 2006).
41. International Energy Agency. CO_2 emissions from fuel combustion statistics, 2016. *Organization for Economic Cooperation and Development (OECD)* https://www.oecd-ilibrary.org/energy/co2-emissions-from-fuel-combustion-2016_co2_fuel-2016-en (2016).
42. National Minerals Information Center. Commodity Statistics and Information. *United States Geological Survey* https://www.usgs.gov/centers/nmic/commodity-statistics-and-information (2016).
43. Davis, S. C. & Diegel, S. W. *Transportation Energy Data Book* 25th edn (Center for Transportation Analysis, Oak Ridge National Laboratory, 2006).
44. Zheng, B. et al. High-resolution mapping of vehicle emissions in China in 2008. *Atmos. Chem. Phys.* **14**, 9787–9805 (2014).
45. Forster, P. et al. *Mitigation Pathways Compatible with 1.5 °C in the Context of Sustainable Development: Supplementary Material* https://report.ipcc.ch/sr15/pdf/sr15_chapter2_supplementary_materials.pdf (2018).
46. Tanaka, K. & O'Neill, B. C. The Paris Agreement zero-emissions goal is not always consistent with the 1.5 °C and 2 °C temperature targets. *Nat. Clim. Chang.* **8**, 319–324 (2018).
47. Millar, R. J. et al. Emission budgets and pathways consistent with limiting warming to 1.5 °C. *Nat. Geosci.* **10**, 741–747 (2017); correction **11**, 454–455 (2018).
48. Jones, C. D. et al. Simulating the Earth system response to negative emissions. *Environ. Res. Lett.* **11**, 095012 (2016).
49. Schurer, A. P., Mann, M. E., Hawkins, E., Tett, S. F. B. & Hegerl, G. C. Importance of the pre-industrial baseline for likelihood of exceeding Paris goals. *Nat. Clim. Chang.* **7**, 563–567 (2017).
50. Lowe, J. A. & Bernie, D. The impact of Earth system feedbacks on carbon budgets and climate response. *Phil. Trans. Royal Soc. A* **376**, 20170263 (2018).
51. Comyn-Platt, E. et al. Carbon budgets for 1.5 and 2 °C targets lowered by natural wetland and permafrost feedbacks. *Nat. Geosci.* **11**, 568–573 (2018); correction **11**, 882–886 (2018).
52. Gasser, T. et al. Path-dependent reductions in CO_2 emission budgets caused by permafrost carbon release. *Nat. Geosci.* **11**, 830–835 (2018); correction **12**, 80 (2019).
53. MacDougall, A. H., Zickfeld, K., Knutti, R. & Matthews, H. D. Sensitivity of carbon budgets to permafrost carbon feedbacks and non-CO_2 forcings. *Environ. Res. Lett.* **10**, 125003 (2015); correction **11**, 019501 (2016).
54. Mengis, N., Partanen, A.-I., Jalbert, J. & Matthews, H. D. 1.5 °C carbon budget dependent on carbon cycle uncertainty and future non-CO2 forcing. *Sci. Rep.* **8**, 5831 (2018).
55. Rogelj, J., Meinshausen, M., Schaeffer, M., Knutti, R. & Riahi, K. Impact of short-lived non-CO_2 mitigation on carbon budgets for stabilizing global warming. *Environ. Res. Lett.* **10**, 075001 (2015).
56. Alvarez, R. A. et al. Assessment of methane emissions from the U.S. oil and gas supply chain. *Science* **361**, 186–188 (2018).
57. Carlson, K. M. et al. Greenhouse gas emissions intensity of global croplands. *Nat. Clim. Chang.* **7**, 63–68 (2017).
58. Hirth, L. & Steckel, J. C. The role of capital costs in decarbonizing the electricity sector. *Environ. Res. Lett.* **11**, 114010 (2016).
59. Meunier, G., Ponssard, J.-P. & Thomas, C. Capacity investment under demand uncertainty: the role of imports in the U.S. cement industry: capacity investment under demand uncertainty. *J. Econ. Manage. Strategy* **25**, 455–486 (2016).
60. U.S. Energy Information Administration. *Levelized Cost and Levelized Avoided Cost of New Generation Resources in the Annual Energy Outlook*. Report No AEO2019 https://www.eia.gov/outlooks/aeo/pdf/electricity_generation.pdf (2019).
61. U.S. Energy Information Administration. *Capital Cost Estimates for Utility Scale Electricity Generating Plants* https://www.eia.gov/analysis/studies/powerplants/capitalcost/pdf/capcost_assumption.pdf (2016).
62. Schröder, A., Kunz, F., Meiss, J., Mendelevitch, R. & Von Hirschhausen, C. *Data Documentation: Current and Prospective Costs of Electricity Generation until 2050*. Report No. 68 (Deutsches Institut für Wirtschaftsforschung, 2013).
63. Energy Technology Systems Analysis Programme. *Industrial Combustion Boilers*. Technology Brief I01. https://www.etsap.org (International Energy Authority, 2010).
64. Energy Technology Systems Analysis Programme. *Cooking Appliances*. Technology Brief R06. https://www.etsap.org (International Energy Authority, 2012).

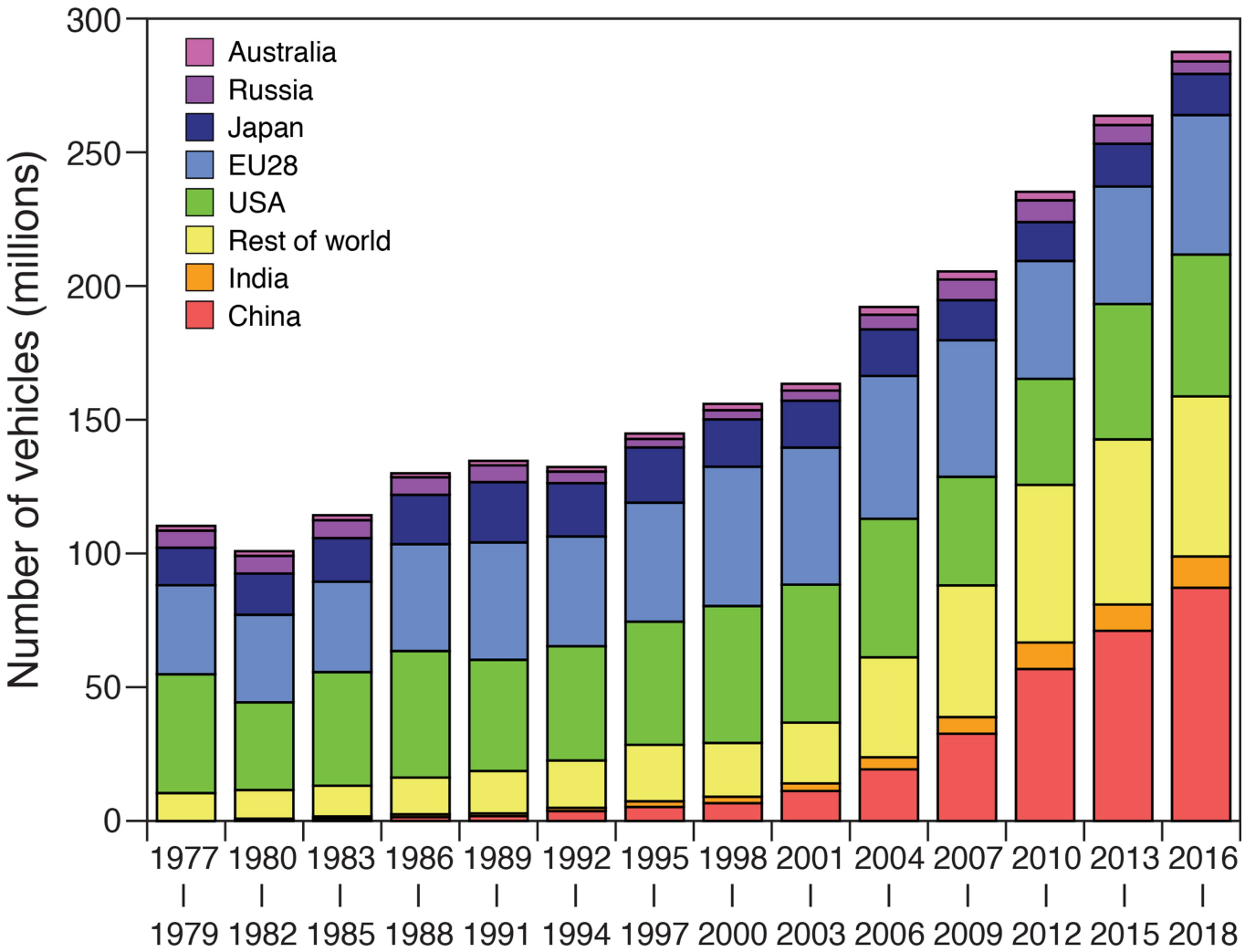
65. International Energy Agency. *Energy Statistics and Balances of OECD Countries 2015*. https://www.iea.org/classicstats/relateddatabases/worldenergystatisticsandbalances/(2016).
66. International Energy Agency. *Energy Statistics and Balances of Non-OECD Countries 2015*, (2016).
67. Storchmann, K. On the depreciation of automobiles: an international comparison. *Transportation* **31**, 371–408 (2004).
68. Liu, F. et al. High-resolution inventory of technologies, activities, and emissions of coal-fired power plants in China from 1990 to 2010. *Atmos. Chem. Phys.* **15**, 18787–18837 (2015).
69. Janssens-Maenhout, G. et al. HTAP_v2.2: a mosaic of regional and global emission grid maps for 2008 and 2010 to study hemispheric transport of air pollution. *Atmos. Chem. Phys.* **15**, 11411–11432 (2015).



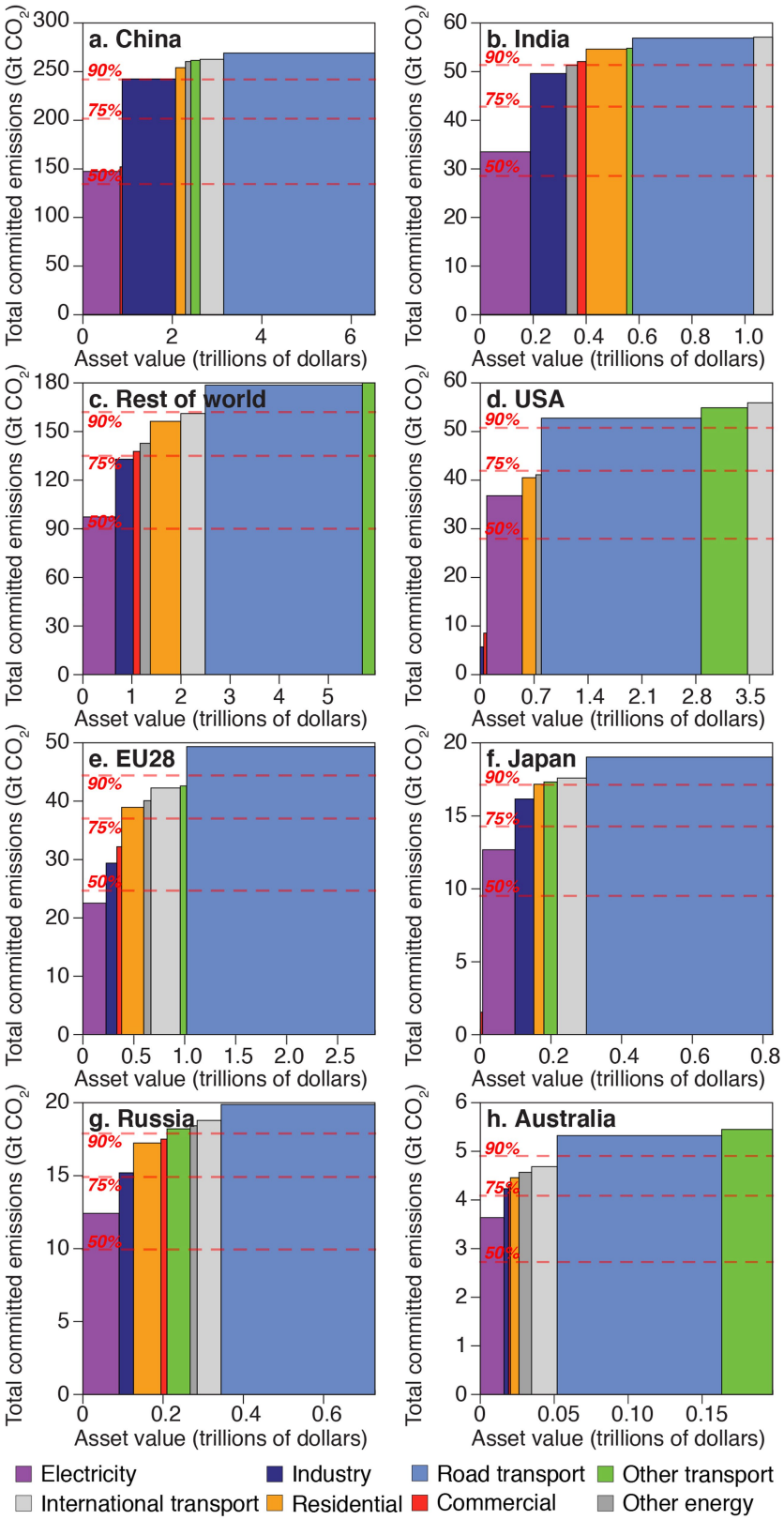
Extended Data Fig. 1 | Changes in commitments from existing energy infrastructure. a, **b**, Estimates of future CO_2 emissions every four years (1998, 2002, 2006, 2010, 2014 and 2018) by industry sector (**a**) and country/region (**b**), assuming historical lifetimes and utilization rates. **c**, **d**, Corresponding changes in remaining commitments by industry sector (**c**) and country/region (**d**).



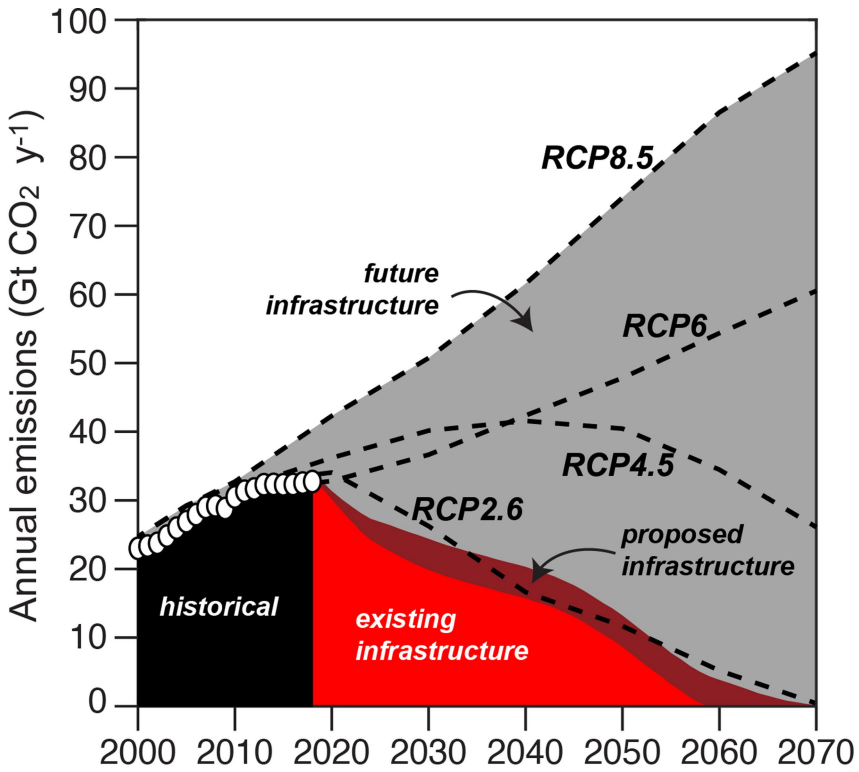
Extended Data Fig. 2 | Age structure of Chinese major industrial capacity. a, **b**, The operating capacity of raw steel in the iron and steel industry (**a**) and clinker in the cement industry (**b**). The youngest units are shown at the bottom.



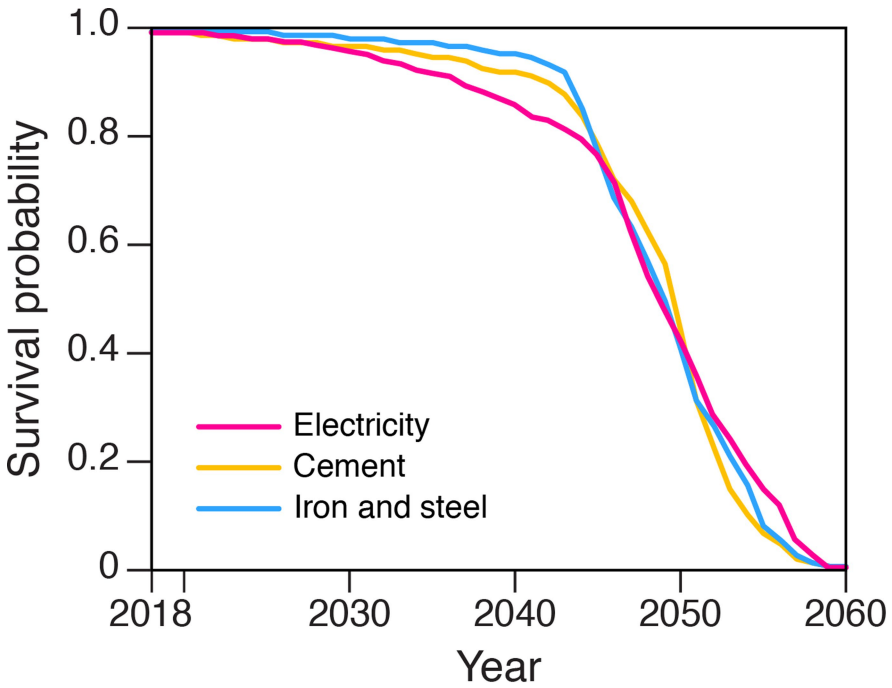
Extended Data Fig. 3 | Age structure of existing road-transport infrastructure. This figure shows the numbers of vehicle sales by country/region.



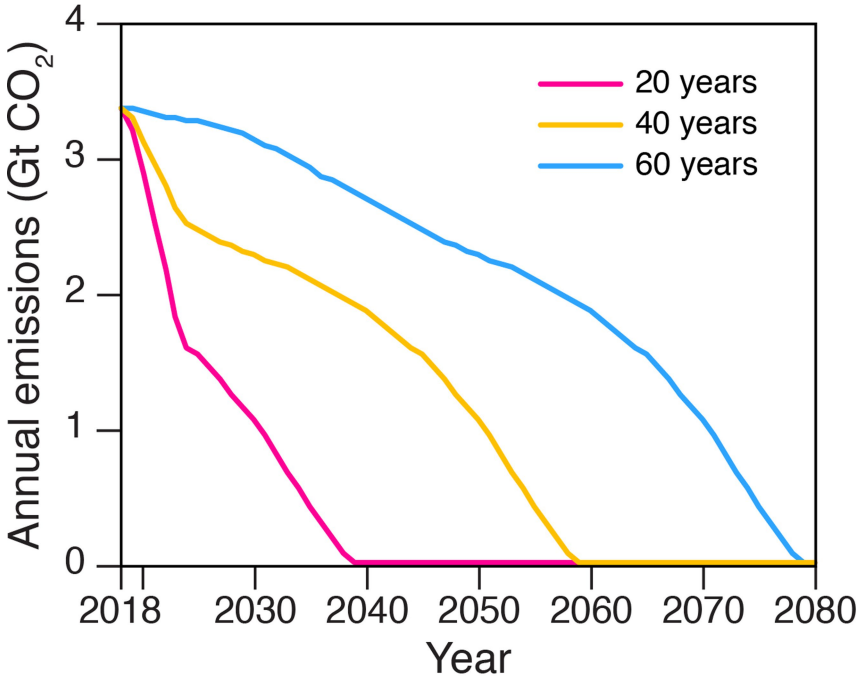
Extended Data Fig. 4 | Asset values and committed emissions for existing infrastructure. Total committed emissions are plotted against asset value, by country/region and sector. Dashed horizontal lines indicate 50%, 75% and 90% of total committed emissions if operated as historically.



Extended Data Fig. 5 | Annual emissions from existing, proposed and future infrastructure. The figure shows historical CO_2 emissions from fossil-fuel energy infrastructure (black), and future CO_2 emissions from existing (red) and proposed (dark red) energy infrastructure, as well as future infrastructure (dark grey) under particular representative concentration pathways (RCPs: RCP8.5, RCP6, RCP4.5 and RCP2.6).



Extended Data Fig. 6 | Survival curves for power and major industries in China. This figure shows survival curves for the electricity sector, cement industry, and iron and steel industry in China under the assumption of 40-year lifetimes.



Extended Data Fig. 7 | Annual emissions from residential, commercial and other energy infrastructure. The figure shows future annual CO_2 emissions from residential, commercial and other energy infrastructure under the assumptions of 20-, 40- and 60-year lifetimes.

Extended Data Table 1 | Comparison of committed emissions

		Ref. 8		Ref. 9		Ref. 11		Ref. 12		Ref. 13		This study	
		Gt CO_2	Year of dataset	Gt CO_2	Year of dataset	Gt CO_2	Year of dataset	Gt CO_2	Year of dataset	Gt CO_2	Year of dataset	Gt CO_2	Year of dataset
Existing	Electricity	224	2009	307	2012	-	-	308	2016	345 (261-451)	2009 *	358 (240-493) †	2018
	Coal		2009	206	2012	190	2016	220	2016	-	-	260 (175-358)	2018
	Gas, oil, and other fuels		2009	100	2012	-	-	88	2016	-	-	98 (65-135)	2018
	Industry	104	2009			-	-	-	-	154 (117-191)	2009	162 (110-219)	2017
	Transport	116	2009			-	-	-	-	92 (73-110)	2017	64 (53-75)	2017
	Residential, commercial, and other energy	53	2009			-	-	-	-	121 (91-158)	2009	74 (52-105)	2018
	All Sectors	**496 (282-701)**				-	-	-	-	**715 (546-909)**	-	**658 (455-892)**	-
Proposed	Electricity					-	-	271	2016	-	-	188 (142-234)	2018
	Coal					150	2016	210	2016	-	-	97 (74-121)	2018
	Gas, oil, and other fuels					-	-	61	2016	-	-	91 (68-113)	2018
All Sectors + Proposed Electricity												**846 (597-1,126)**	

Comparison of committed emissions by sector, estimated here and previously[8,9,11–13]. Note that, in some cases, the totals may not correspond to the sum of the underlying sectors, owing to rounding.
*The age distribution of infrastructure is the same as in 2009, but annual emissions from the infrastructure were adjusted up to 2018 levels.
†The range represents the committed emissions estimated under the assumption of 30–50-year lifetimes for all sectors except transportation (12–18-year lifetimes).

LETTER

https://doi.org/10.1038/s41586-019-1462-2

Metamorphism and the evolution of plate tectonics

Robert M. Holder[1,2]*, Daniel R. Viete[1], Michael Brown[3] & Tim E. Johnson[4,5]

Earth's mantle convection, which facilitates planetary heat loss, is manifested at the surface as present-day plate tectonics[1]. When plate tectonics emerged and how it has evolved through time are two of the most fundamental and challenging questions in Earth science[1–4]. Metamorphic rocks—rocks that have experienced solid-state mineral transformations due to changes in pressure (*P*) and temperature (*T*)—record periods of burial, heating, exhumation and cooling that reflect the tectonic environments in which they formed[5,6]. Changes in the global distribution of metamorphic (*P*, *T*) conditions in the continental crust through time might therefore reflect the secular evolution of Earth's tectonic processes. On modern Earth, convergent plate margins are characterized by metamorphic rocks that show a bimodal distribution of apparent thermal gradients (temperature change with depth; parameterized here as metamorphic *T*/*P*) in the form of paired metamorphic belts[5], which is attributed to metamorphism near (low *T*/*P*) and away from (high *T*/*P*) subduction zones[5,6]. Here we show that Earth's modern plate tectonic regime has developed gradually with secular cooling of the mantle since the Neoarchaean era, 2.5 billion years ago. We evaluate the emergence of bimodal metamorphism (as a proxy for secular change in plate tectonics) using a statistical evaluation of the distributions of metamorphic *T*/*P* through time. We find that the distribution of metamorphic *T*/*P* has gradually become wider and more distinctly bimodal from the Neoarchaean era to the present day, and the average metamorphic *T*/*P* has decreased since the Palaeoproterozoic era. Our results contrast with studies that inferred an abrupt transition in tectonic style in the Neoproterozoic era (about 0.7 billion years ago[1,7,8]) or that suggested that modern plate tectonics has operated since the Palaeoproterozoic era (about two billion years ago[9–12]) at the latest.

The theory of plate tectonics can explain the assembly and break-up of supercontinents, how mountain ranges and major mineral deposits form[13], and perhaps even why there is life on Earth[14,15]. However, it is unclear how plate tectonics emerged and why it does not occur on other planets in the Solar System. Although there is broad agreement that plate tectonics has been dominant during the last billion years of our planet's history, how and when it emerged, and how it has evolved through time, are disputed[2,3,6,7]. A diagnostic feature of plate tectonics on modern Earth is the bimodal distribution of metamorphic temperatures and pressures (Fig. 1), which is expressed in paired metamorphic belts[5,6] and is key to the identification of the past operation of plate tectonics from the rock record. Here we present a statistical evaluation of metamorphic *T*/*P* through Earth's history, with the purpose of documenting the emergence and evolution of the bimodal distribution of metamorphic *T*/*P* as a proxy for the emergence and evolution of Earth's plate tectonic regime. We show that the modern bimodal distribution of metamorphic *T*/*P*, and therefore modern plate tectonics, developed gradually since the end of the Neoarchaean era, about 2.5 Gyr ago. We hypothesize that the development of modern plate tectonics is linked to secular cooling of the mantle and associated changes in the thickness, buoyancy and rheology of oceanic lithosphere, resulting in evolution in the styles of both subduction and collisional orogenesis.

We assess changes in the global distribution of metamorphic *T*/*P* using a database of age and (*P*, *T*) conditions of metamorphic rocks from 564 localities from ref. [4]. The data define distinct modes centred at about 2.6, 1.8, 1.0, 0.5 and <0.2 Gyr ago[4] (Extended Data Fig. 1).

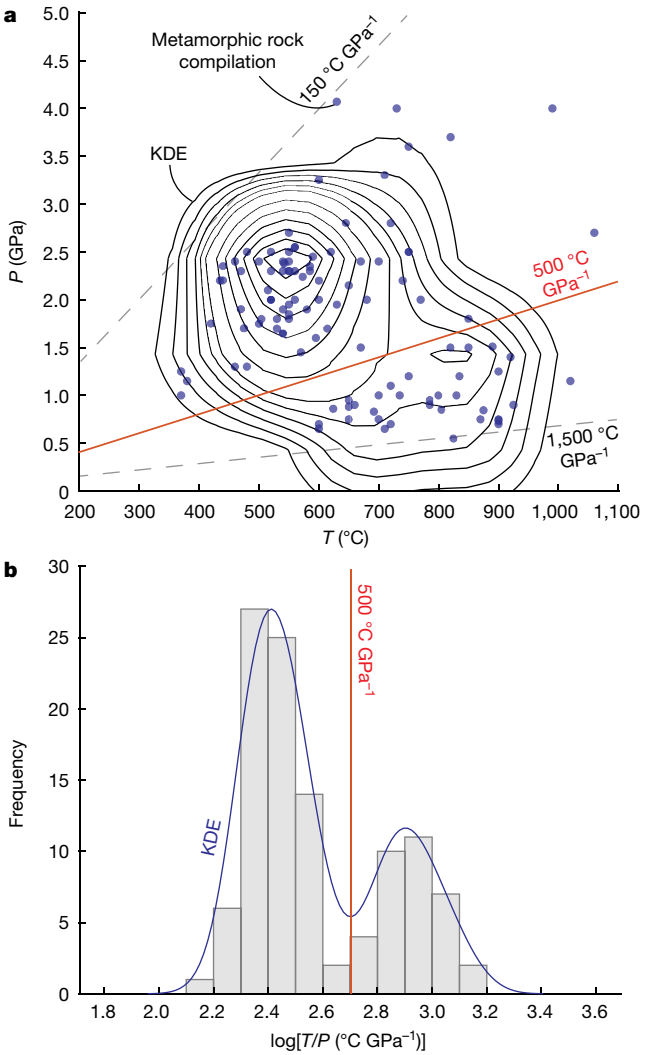


Fig. 1 | Metamorphism in the last 0.2 Gyr is characterized by a bimodal distribution of apparent metamorphic thermal gradients, *T*/*P*. a, Kernel density estimates (KDE) of metamorphic (*P*, *T*) conditions in rocks younger than 0.2 Gyr (ref. [4]). The red line (500 °C GPa^{-1}) represents the dividing line between the two modes shown in **b**. **b**, Histogram and KDE of metamorphic *T*/*P*, plotted logarithmically to illustrate more clearly the bimodal distribution shown in **a**. This bimodal distribution of *T*/*P* is manifest in paired metamorphic belts and is considered a diagnostic feature of plate tectonics.

[1]Morton K. Blaustein Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, MD, USA. [2]Department of Earth and Environmental Sciences, University of Michigan, Ann Arbor, MI, USA. [3]Laboratory for Crustal Petrology, Department of Geology, University of Maryland, College Park, MD, USA. [4]School of Earth and Planetary Sciences, The Institute for Geoscience Research (TIGeR), Space Science and Technology Centre, Curtin University, Perth, Western Australia, Australia. [5]State Key Laboratory of Geological Processes and Mineral Resources, China University of Geosciences, Wuhan, China. *e-mail: roholder@umich.edu

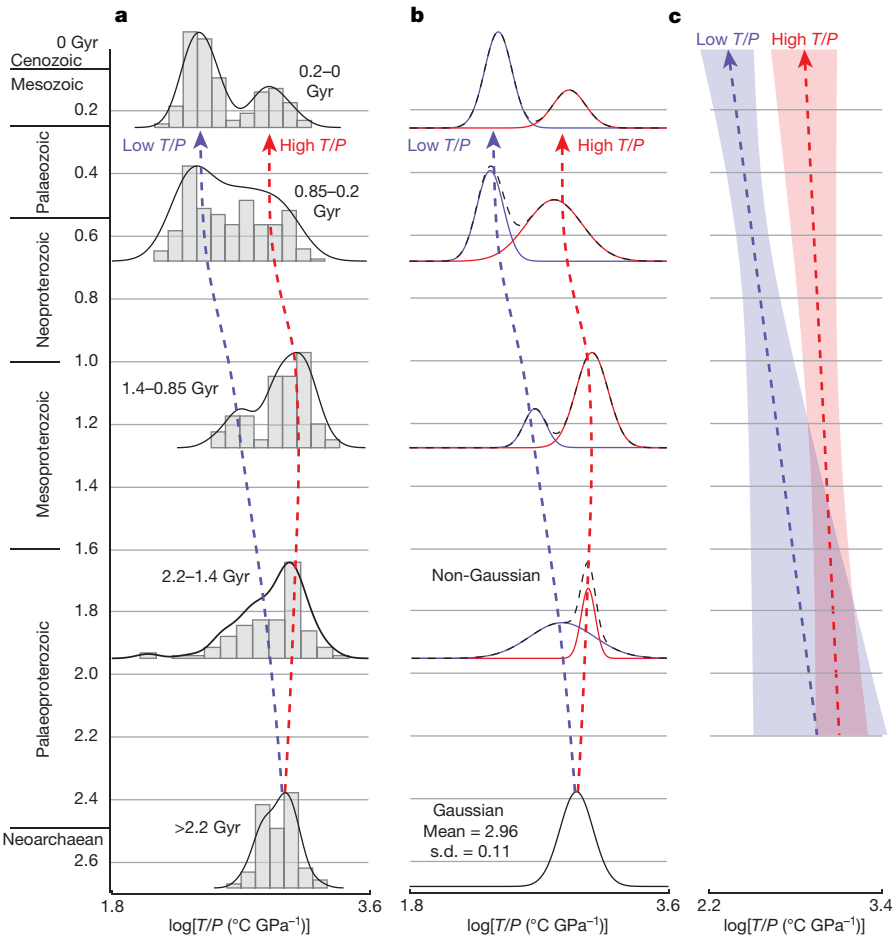


Fig. 2 | The bimodal distribution of modern metamorphism evolved gradually since the end of the Neoarchaean era. a, Histograms and KDEs of metamorphic *T/P* since the Neoarchaean era. **b**, Fits of the distributions of metamorphic *T/P* shown in **a** with bimodal Gaussian mixing models. Global metamorphism >2.2 Gyr ago can be fitted by a unimodal Gaussian distribution (95% confidence interval, $n = 72$, p value of 0.23); metamorphism ≤2.2 Gyr ago is non-Gaussian (95% confidence intervals; 2.2–1.4 Gyr ago, $n = 106$, $p = 5.4 \times 10^{-5}$; 1.4–0.85 Gyr ago, $n = 45$, $p = 0.017$; 0.85–0.2 Gyr ago, $n = 232$, $p = 3.6 \times 10^{-7}$; 0.2–0 Gyr ago, $n = 109$, $p = 2.2 \times 10^{-7}$), but is well described by bimodal mixed-Gaussian distributions. **c**, Linear regressions of low-*T/P* and high-*T/P* modes (from **b**) with 95% confidence envelopes. The best-fit bimodal distributions of metamorphic *T/P* become increasingly distinct, showing divergence and cooling of the low-*T/P* and high-*T/P* modes through time.

These peaks broadly correspond to peaks in the global distribution of igneous and detrital zircon U–Pb ages[16–18] and, similarly to those studies, we interpret the peaks to reflect cyclicity[4]. To provide more statistically robust (larger number of data points per calculation) characterization of secular changes in the distributions of metamorphic *T/P* from one cycle to the next, we binned the data about each of these peaks, as shown in Extended Data Fig. 1. Histograms and kernel density estimates (KDE) of each distribution are shown in Fig. 2a. With decreasing age: (1) the KDE of >2.2-Gyr-old metamorphism is narrow and symmetric; (2) the KDE of 2.2–1.4-Gyr-old metamorphism is skewed towards lower *T/P*, with three low-*T/P* eclogite outliers[11,12,19]; (3) the KDE of 1.4–0.85-Gyr-old metamorphism shows a distinct low-*T/P* mode that is not apparent in the >1.4-Gyr-old KDEs; (4) the KDE of 0.85–0.2-Gyr-old metamorphism is notably broader than the older distributions, with a prominent mode at lower *T/P*; and (5) the KDE of <0.2-Gyr-old ('modern') metamorphism is bimodal, with distinct high-*T/P* and low-*T/P* peaks that are much more pronounced than for the older distributions. Qualitatively, the secular evolution in the KDEs is interpreted to represent a gradual transition from a narrow, unimodal distribution of metamorphic *T/P* in the Archaean eon, to a distinctly bimodal distribution in the modern metamorphic rock record.

To assess this interpretation quantitatively, we fitted a Gaussian mixing model to each distribution (Fig. 2b). First, the distributions were assessed by whether they are non-Gaussian (that is, do not represent a single Gaussian distribution) at a 95% confidence interval, according to the Shapiro–Wilk test[20]. If the data distributions were assessed to be non-Gaussian, we fitted a bimodal mixed-Gaussian distribution (that is, two Gaussian distributions; Fig. 2b, Extended Data Table 1) to evaluate how these compare to the bimodal distribution of the modern metamorphic rock record (Figs. 1, 2a). The distribution of data older than 2.2 Gyr is Gaussian with a mean and standard deviation of 2.96 and 0.11 (for log[*T/P* (°C GPa^{-1})]), respectively. By contrast, each of the younger (<2.2 Gyr old) distributions is non-Gaussian but is well described by a bimodal mixed-Gaussian distribution. The difference between the two best-fit Gaussian distributions ('low-*T/P*' and 'high-*T/P*' in Fig. 2b) increases through time (Fig. 2c). Viewed together, the KDEs and modelled mixed-Gaussian distributions show a continuous increase in the variability of thermal gradients recorded by metamorphic rocks, as well as a gradual emergence of a discrete and prominent low-*T/P* mode of metamorphism since the end of the Archaean eon.

The Palaeoproterozoic era (about 2 Gyr ago) is notable in that it has three distinct outliers with low-*T/P* gradients comparable to those of rocks formed in modern cold collisional environments[11,12,19] (Fig. 3). These data have been used to suggest that subduction and collision similar to those of modern Earth may have been operative during the Palaeoproterozoic era[11,12,19]. Alternatively, the data may reflect local, anomalous subduction–collision similar to that in modern tectonic environments, but not representative of the dominant

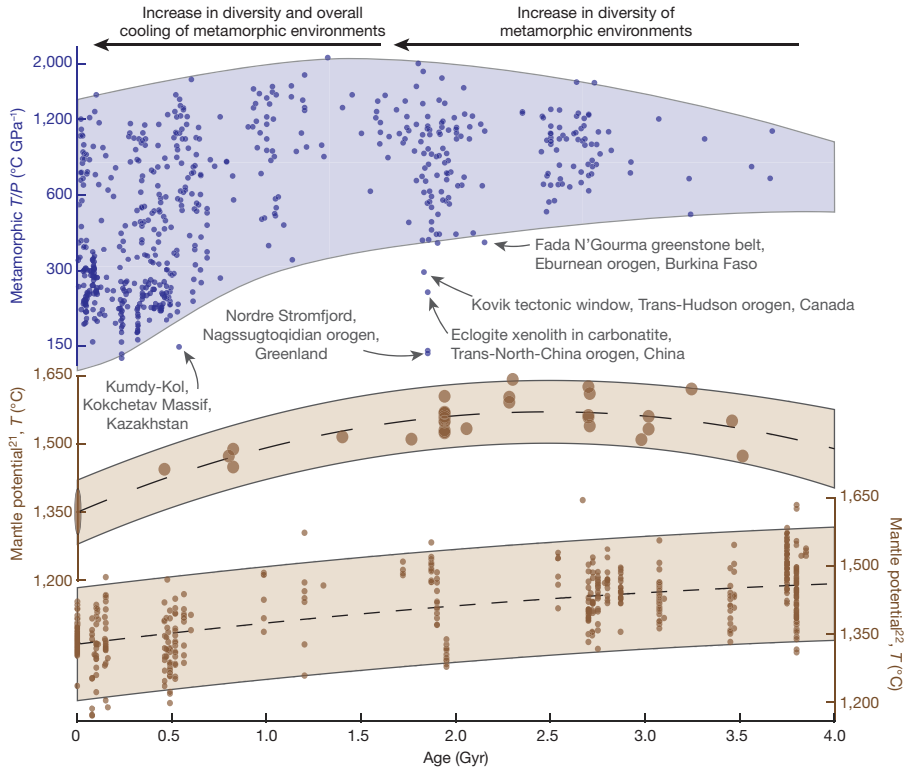


Fig. 3 | The range of metamorphic *T/P* (blue symbols) has become increasingly varied through time, with its average value decreasing since about 2 Gyr ago. Metamorphic rocks that fall outside this trend (anomalously low *T/P* for their age) are labelled. The overall lowering of metamorphic *T/P* (and increase in metamorphic diversity) through time coincides with secular cooling of the upper mantle. Large brown symbols are from ref. [21], with ages modified as described in ref. [30]. Small brown symbols are from ref. [22], modified for direct comparison with ref. [21], as described in Methods. Both datasets are shown with best-fit quadratic regressions and 95% prediction intervals.

global tectonic regime during the Palaeoproterozoic era[4]. Here we focus on the broader, continuous global trends in metamorphic *T/P* outlined above.

From the statistical evaluation of the distributions of metamorphic *T/P* through time presented here, it is hypothesized that the modern style of plate tectonics—characterized by a distinctly bimodal distribution of metamorphic *T/P*—developed gradually (Fig. 3). This hypothesis can be considered an alternative to: (1) hypotheses that infer plate tectonics to have begun abruptly in the Neoproterozoic era, based on the apparently sudden appearance of blueschist and ultrahigh-pressure (UHP) metamorphism about 0.7 Gyr ago[7]; (2) hypotheses that infer that a plate tectonic regime similar to the modern one has been operative since the Palaeoproterozoic era (about 2 Gyr ago)[10–12] at the latest, and; (3) the hypothesis of transient (≤0.3 Gyr) modern-like plate tectonic behaviour in the Palaeoproterozoic era, before true plate tectonics began in the Neoproterozoic era[1]. The hypothesis presented here, of a gradual but continuous transition in tectonic style since about 2.5 Gyr ago, is similar to the hypothesized gradual onset of plate tectonics between 3.2 and 2.5 Gyr ago proposed in ref. [2], but extends that gradual change in tectonic style through to the modern era. We argue that the most plausible mechanism for the hypothesized gradual change in plate tectonic style since the Archaean eon is secular cooling of the upper mantle (Fig. 3).

Changes in the temperature of the upper mantle affect not only the thermal state of the crust but also the thickness and density of oceanic lithosphere. Under the higher upper-mantle temperatures that are thought to have existed during the Proterozoic and Archaean eons[21–23], the degree of decompression melting at mid-ocean ridges is predicted to have been higher, resulting in thicker oceanic lithosphere[23,24] (as much as 135 km 2.5 Gyr ago versus 60 km today)[25], which could have remained buoyant (relative to the underlying asthenosphere) for substantially longer than modern oceanic lithosphere (as much as 0.1 Gyr to attain neutral buoyancy 2.5 Gyr ago versus 0.01–0.03 Gyr today)[25]. Greater thickness and buoyancy of oceanic lithosphere before 1 Gyr ago might have favoured more uniformly shallow (less steeply dipping) subduction and overall higher thermal gradients in subduction environments, similar to the moderate-*T/P* metamorphism (greenschist–amphibolite–eclogite series) associated with modern collisional orogenesis.

A possible young analogue for shallower, hotter subduction metamorphism that might have been more prevalent before 1 Gyr ago is the greenschist–amphibolite-facies Orocopia–Pelona–Rand schist (OPRS) of southern California, which records *T/P* (500–650 °C GPa^{-1})[26] comparable to many eclogites, high-pressure granulites and amphibolites in the Palaeo- and Mesoproterozoic eras (2.5–1.0 Gyr ago; Fig. 3, Extended Data Fig. 2). The OPRS is thought to have formed in response to a transition from steeper, colder subduction (Franciscan-type) to shallower, hotter subduction related to the incoming of an oceanic plateau (thicker, more buoyant oceanic lithosphere)[26]. Consideration of similar >1-Gyr-old rocks as plausibly related to subduction, rather than focusing only on blueschist and UHP metamorphism, might offer new insights into subduction processes (oceanic and continental) on the early Earth. Many 2.5–1.0-Gyr-old orogenic belts—including the Grenville, Sveconorwegian, Trans-North China, Trans-Hudson, Eburnean, Ubendian–Usagaran and Belomorian belts—preserve bimodal distributions of metamorphic rocks, with the lowest-*T/P* rocks characterized by *T/P* similar to that of the OPRS (about 500–650 °C GPa^{-1}; Extended Data Fig. 2)[26]. These contrast the blueschist and UHP metamorphism that are common on modern Earth. We hypothesize that modern-style plate tectonics might have developed as the time to neutral buoyancy became substantially less than the average age of oceanic lithosphere, favouring colder, steeper subduction of the type most common in modern subduction zones.

This study focuses on metamorphic *T/P*—a proxy for the thermal gradients of different tectonic environments—as a more reliable indicator of the tectonic regime than either *T* or *P*. However, another

important secular change in metamorphism relates to a large increase in the maximum pressure of metamorphism through the Proterozoic eon (2.5–0.55 Gyr ago), resulting in the widespread occurrence of continental UHP metamorphism beginning in the Neoproterozoic era, about 0.7 Gyr ago[4]. The *T/P* values of UHP rocks are consistent with the gradual change proposed here (Figs. 2, 3), but the reason for their higher maximum pressures requires additional comment. Although many mechanisms have been invoked to explain the formation and exhumation of continental UHP rocks[27], their observation at Earth's surface in general requires that: (1) positively buoyant material be carried to depth; and (2) the downward force acting on the material subsequently ceases (for example, owing to physical separation from the denser material driving subduction or foundering), allowing the material to return towards Earth's surface. The geodynamic controls on the formation of UHP rocks have been addressed through numerical experiments by ref. [28], which hypothesized that the formation and exhumation of UHP rocks might have been precluded by shallower slab breakoff associated with a hotter mantle (and rheologically weaker plates) earlier in Earth's history. Although further simulations are needed to understand slab breakoff during subduction, these results suggest that the increase in maximum metamorphic pressure, the emergence of a distinct low-*T/P* mode of metamorphism (Fig. 2) and the overall decrease in metamorphic *T/P* since the end of the Neoarchaean era, 2.5 Gyr ago (Fig. 2), can be all linked to gradual cooling of the upper mantle (Fig. 3).

In summary, we present a statistical evaluation of metamorphic *T/P* through Earth's history, with the purpose of documenting the gradual emergence of the modern bimodal distribution of metamorphic *T/P* as a proxy for the emergence of Earth's modern plate tectonic regime. This approach is rooted in the classical concept of paired metamorphic belts[5] as a diagnostic feature of plate tectonics, while leaving open the possibility that, owing to evolution in the style of plate tectonics[29], paired metamorphism in Earth's past might have been characterized by different apparent thermal gradients than today. We show that the modern bimodal distribution of metamorphic *T/P* developed gradually since the end of the Neoarchaean era, about 2.5 Gyr ago, and that globally bimodal metamorphism emerged before either blueschist or UHP metamorphism in the geological record. We hypothesize that both the development of bimodal metamorphism and the appearance of blueschist and UHP metamorphism are linked to secular cooling of the mantle and to associated changes in the thickness, buoyancy and rheology of oceanic lithosphere, resulting in evolution in the styles of both subduction and collisional orogenesis. Importantly, this and other hypotheses concerning changes in tectonic style (or lack thereof)[1,4,6,7,10–12,19] through time are testable through further numerical modelling and continued examination of the Precambrian metamorphic rock record.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1462-2.

Received: 18 January 2019; Accepted: 30 May 2019;
Published online 7 August 2019.

1. Stern, R. J. The evolution of plate tectonics. *Philos. Trans. R. Soc. A* **376**, 20170406 (2018).
2. Cawood, P. A. et al. Geological archive of the onset of plate tectonics. *Philos. Trans. R. Soc. A* **376**, 20170405 (2018).
3. Korenaga, J. Crustal evolution and mantle dynamics through Earth history. *Philos. Trans. R. Soc. A* **376**, 20170408 (2018).
4. Brown, M. & Johnson, T. Time's arrow, time's cycle: granulite metamorphism and geodynamics. *Mineral. Mag.* **83**, 323–338 (2019).
5. Miyashiro, A. Evolution of metamorphic belts. *J. Petrol.* **2**, 277–311 (1961).
6. Brown, M. Duality of thermal regimes is the distinctive characteristic of plate tectonics since the Neoarchean. *Geology* **34**, 961–964 (2006).
7. Stern, R. J. Evidence from ophiolites, blueschists, and ultrahigh-pressure metamorphic terranes that the modern episode of subduction tectonics began in Neoproterozoic time. *Geology* **33**, 557–560 (2005).
8. Stern, R. J., Leybourne, M. I. & Tsujimori, T. Kimberlites and the start of plate tectonics. *Geology* **44**, 799–802 (2016).
9. Harrison, T. M. The Hadean crust: evidence from >4 Ga zircons. *Annu. Rev. Earth Planet. Sci.* **37**, 479–505 (2009).
10. Ganne, J. et al. Modern-style plate subduction preserved in the Palaeoproterozoic West African craton. *Nat. Geosci.* **5**, 60–65 (2012).
11. Weller, O. M. & St-Onge, M. R. Record of modern-style plate tectonics in the Palaeoproterozoic Trans-Hudson orogen. *Nat. Geosci.* **10**, 305–311 (2017).
12. Xu, C. et al. Cold deep subduction recorded by remnants of a Paleoproterozoic carbonated slab. *Nat. Commun.* **9**, 1–8 (2018).
13. Goldfarb, R. J., Bradley, D. & Leach, D. L. Secular variation in economic geology. *Econ. Geol.* **105**, 459–465 (2010).
14. Russell, M. J., Hall, A. J. & Martin, W. Serpentinization as a source of energy at the origin of life. *Geobiology* **8**, 355–371 (2010).
15. Turner, S., Rushmer, T., Reagan, M. & Moyen, J. F. Heading down early on? Start of subduction on Earth. *Geology* **42**, 139–142 (2014).
16. Condie, K. C. & Aster, R. C. Episodic zircon age spectra of orogenic granitoids: the supercontinent connection and continental growth. *Precambr. Res.* **180**, 227–236 (2010).
17. Puetz, S. J., Ganade, C. E., Zimmermann, U. & Borchardt, G. Statistical analyses of Global U-Pb Database 2017. *Geoscience Frontiers* **9**, 121–145 (2018).
18. Campbell, I. H. & Allen, C. M. Formation of supercontinents linked to increases in atmospheric oxygen. *Nat. Geosci.* **1**, 554–558 (2008).
19. Glassley, W. E., Korstgard, J. a., Sorensen, K. & Platou, S. W. A new UHP metamorphic complex in the 1.8 Ga Nagssugtoqidian Orogen of West Greenland. *Am. Mineral.* **99**, 1315–1334 (2014).
20. Shapiro, S. S. & Wilk, M. B. An analysis of variance test for normality (complete samples). *Biometrika* **52**, 591–611 (1965).
21. Herzberg, C., Condie, K. & Korenaga, J. Thermal history of the Earth and its petrological expression. *Earth Planet. Sci. Lett.* **292**, 79–88 (2010).
22. Condie, K. C., Aster, R. C. & Van Hunen, J. A great thermal divergence in the mantle beginning 2.5 Ga: geochemical constraints from greenstone basalts and komatiites. *Geoscience Frontiers* **7**, 543–553 (2016).
23. Keller, B. & Schoene, B. Plate tectonics and continental basaltic geochemistry throughout Earth history. *Earth Planet. Sci. Lett.* **481**, 290–304 (2018).
24. Langmuir, C. H., Klein, E. M. & Plank, T. in *Mantle Flow and Melt Generation at Mid-Ocean Ridges* (eds Morgan, J. P. et al.) 183–280 (American Geophysical Union, 1992).
25. Korenaga, J. in *Archean Geodynamics and Environments* (eds Benn, K. et al.) 7–32 (American Geophysical Union, 2006).
26. Chapman, A. D. The Pelona–Orocopia–Rand and related schists of southern California: a review of the best-known archive of shallow subduction on the planet. *Int. Geol. Rev.* **59**, 664–701 (2017).
27. Hacker, B. R. & Gerya, T. V. Paradigms, new and old, for ultrahigh-pressure tectonism. *Tectonophysics* **603**, 79–88 (2013).
28. Sizova, E., Gerya, T. & Brown, M. Contrasting styles of Phanerozoic and Precambrian continental collision. *Gondwana Res.* **25**, 522–545 (2014).
29. Chowdhury, P., Gerya, T. & Chakraborty, S. Emergence of silicic continents as the lower crust peels off on a hot plate-tectonic Earth. *Nat. Geosci.* **10**, 698–703 (2017).
30. Johnson, T. E., Brown, M., Kaus, B. J. P. & Vantongeren, J. A. Delamination and recycling of Archaean crust caused by gravitational instabilities. *Nat. Geosci.* **7**, 47–52 (2014).

METHODS

KDEs (Figs. 1, 2a) and bimodal mixed-Gaussian distributions (Fig. 2b, Extended Data Table 1) were calculated using the *ksdensity* and *fitgmdist* functions, respectively, of Matlab release R2018b. The linear regressions and corresponding 95% confidence envelopes, shown in Fig. 2c, were calculated by Monte Carlo analysis (number of regressions: 10,000). For each regression, values of *T*/*P* and age were selected at random from each modelled Gaussian distribution (Fig. 2b, Extended Data Table 1) and the age ranges used to bin the data for the mixed-Gaussian fitting (0–0.2 Gyr, 0.2–0.85 Gyr, 0.85–1.4 Gyr and 1.4–2.2 Gyr). To plot mantle potential temperatures (temperature corrected adiabatically for pressure) from ref. [22] in Fig. 3, the reported 'magma-generation' (*P*, *T*) conditions for 'depleted mantle' compositions were used and corrected for pressure assuming a liquid adiabat of 30 °C GPa^{-1}. Quadratic regressions and 95% prediction envelopes of the mantle potential temperature (Fig. 3) were calculated using the curve-fitting toolbox of Matlab release R2018b.

Data availability

The metamorphic pressure and temperature data used in this study are available online with their original publication (ref. [4]) and in the EarthChem community data repository (http://www.earthchem.org/library; doi:10.1594/IEDA/111316).

Acknowledgements This work was funded by the Morton K. Blaustein Department of Earth and Planetary Sciences at Johns Hopkins University. T.E.J. acknowledges support from the State Key Laboratory for Geological Processes and Mineral Resources, China University of Geosciences, Wuhan (Open Fund GPMR210704). R. Rudnick and B. Hacker provided helpful discussion in the project's infancy. The authors thank P. Cawood and R. Stern for their constructive reviews of this work.

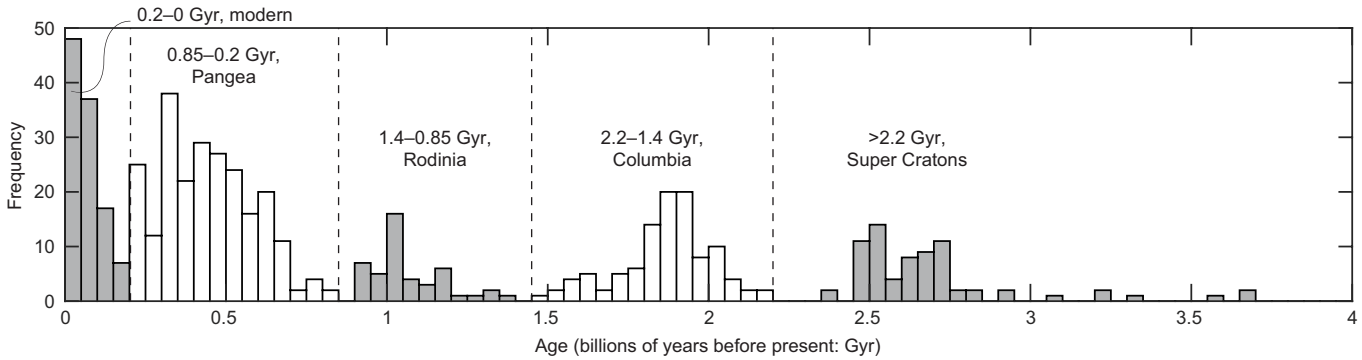
Author contributions R.M.H.: conceptualization, formal analysis, methodology, visualization and writing (original draft, review and editing). D.R.V.: funding acquisition, visualization and writing (original draft, review and editing). M.B.: data curation, investigation, visualization and writing (original draft, review and editing). T.E.J.: data curation, investigation, visualization and writing (original draft, review and editing).

Competing interests The authors declare no competing interests.

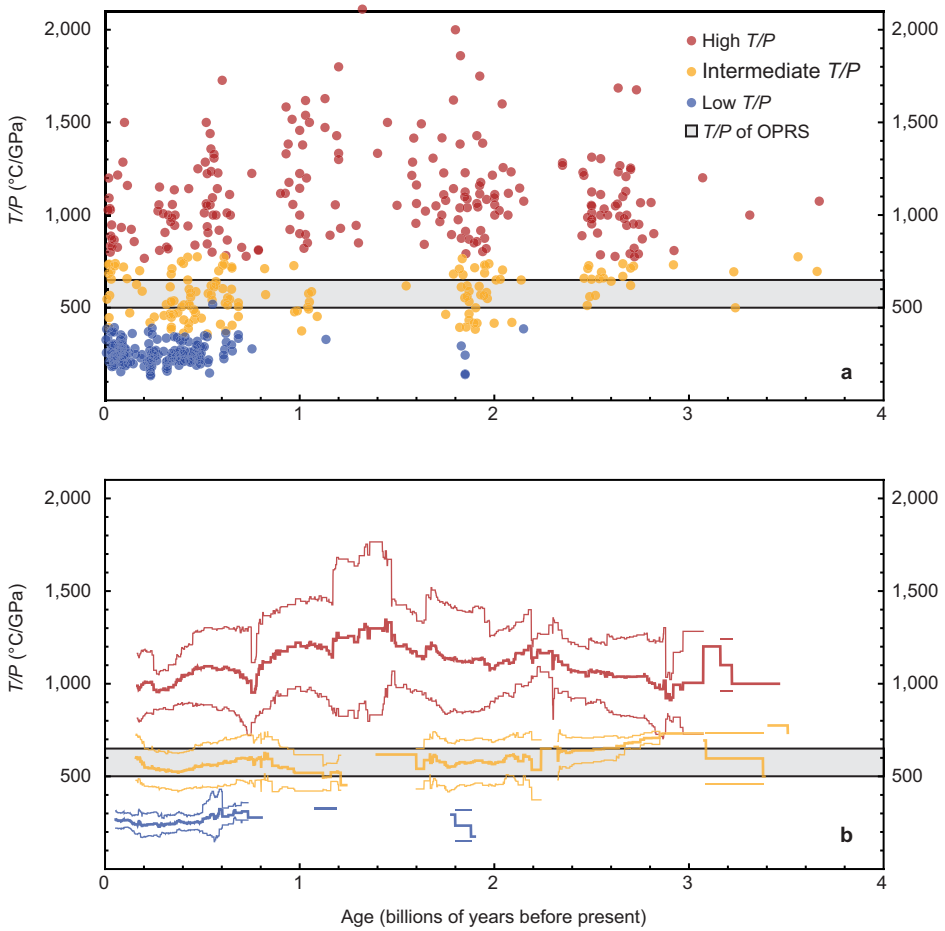
Additional information

Correspondence and requests for materials should be addressed to R.M.H.

Peer review information *Nature* thanks Peter Cawood and Robert Stern for their contribution to the peer review of this work.



Extended Data Fig. 1 | Distribution of ages of metamorphism. The distribution of ages of metamorphism is characterized by peaks at about 2.6, 1.9, 1.0, 0.5 and <0.2 Gyr ago. For the statistical evaluation of metamorphic data presented in this paper, the data were binned about each of these discrete peaks to provide more statistically robust (with higher number of data points) interpretations.



Extended Data Fig. 2 | Comparison between *T/P* values for the Orocopia–Pelona–Rand schist and for the entire dataset used in this study. a, All data are divided into low-, intermediate- and high-*T/P* after ref. [4]. **b**, Moving averages (300-Myr window) and one-standard-deviation envelopes of the data shown in **a**. The OPRS is thought to have formed in response to a transition from steeper, colder subduction ('Franciscan-type') to shallower (more gently dipping), hotter subduction related to the incoming of an oceanic plateau (thicker, more buoyant oceanic lithosphere)[26]. Many Mesoproterozoic and Palaeoproterozoic orogenic belts preserve bimodal distributions of metamorphism, with the lower-*T/P* rocks ('intermediate-*T/P*' in this figure) being characterized by average *T/P* similar to that of the OPRS (about 500–650 °C GPa^{-1})[26], including the Grenville, Sveconorwegian, Trans-North China, Trans-Hudson, Eburnean, Ubendian–Usagaran and Belomorian belts.

Extended Data Table 1 | Results of mixed-Gaussian models

Age (Gyr)	High *T/P*			Low *T/P*		
	Mean (log10 °C/GPa)	s.d.	Component Proportion (%)	Mean (log10 °C/GPa)	s.d.	Component Proportion (%)
0.2–0	2.92	0.11	32	2.42	0.10	68
0.85–0.2	2.81	0.19	59	2.36	0.09	41
1.4–0.85	3.08	0.11	78	2.68	0.08	22
2.2–1.4	3.05	0.05	30	2.86	0.23	70
>2.2	2.96	0.11	100	—	—	—

Mean and standard deviation (s.d.) of the bimodal mixed-Gaussian models shown in Fig. 2b. The distribution of data older than 2.2 Gyr could not be distinguished from a Gaussian distribution at the 95% confidence interval by the Shapiro–Wilk test ($n = 72$, $p = 0.23$); for these data the mean and standard deviation of the best-fit single Gaussian distribution are shown. Each distribution <2.2 Gyr old is distinguishable from a single Gaussian distribution at the 95% confidence interval (2.2–1.4 Gyr old, $n = 106$, $p = 5.4 \times 10^{-5}$; 1.4–0.85 Gyr old, $n = 45$, $p = 0.017$; 0.85–0.2 Gyr old, $n = 232$, $p = 3.6 \times 10^{-7}$; 0.2–0 Gyr ago, $n = 109$, $p = 2.2 \times 10^{-7}$).

LETTER

https://doi.org/10.1038/s41586-019-1440-8

Inhibition of bacterial ubiquitin ligases by SidJ–calmodulin catalysed glutamylation

Sagar Bhogaraju[1,2,3]*, Florian Bonn[1], Rukmini Mukherjee[1,2], Michael Adams[3], Moritz M. Pfleiderer[3], Wojciech P. Galej[3], Vigor Matkovic[1,2], Jaime Lopez-Mosqueda[1,4], Sissy Kalayil[1,2], Donghyuk Shin[1,2,5] & Ivan Dikic[1,2,5]*

The family of bacterial SidE enzymes catalyses phosphoribosyl-linked serine ubiquitination and promotes infectivity of *Legionella pneumophila*, a pathogenic bacteria that causes Legionnaires' disease[1–3]. SidE enzymes share the genetic locus with the *Legionella* effector SidJ that spatiotemporally opposes the toxicity of these enzymes in yeast and mammalian cells, through a mechanism that is currently unknown[4–6]. Deletion of SidJ leads to a substantial defect in the growth of *Legionella* in both its natural hosts (amoebae) and in mouse macrophages[4,5]. Here we demonstrate that SidJ is a glutamylase that modifies the catalytic glutamate in the mono-ADP ribosyl transferase domain of the SdeA, thus blocking the ubiquitin ligase activity of SdeA. The glutamylation activity of SidJ requires interaction with the eukaryotic-specific co-factor calmodulin, and can be regulated by intracellular changes in Ca^{2+} concentrations. The cryo-electron microscopy structure of SidJ in complex with human apo-calmodulin revealed the architecture of this heterodimeric glutamylase. We show that, in cells infected with *L. pneumophila*, SidJ mediates the glutamylation of SidE enzymes on the surface of vacuoles that contain *Legionella*. We used quantitative proteomics to uncover multiple host proteins as putative targets of SidJ-mediated glutamylation. Our study reveals the mechanism by which SidE ligases are inhibited by a SidJ–calmodulin glutamylase, and opens avenues for exploring an understudied protein modification (glutamylation) in eukaryotes.

L. pneumophila contains about 300 effector proteins that modulate the cellular processes of the host through diverse activities, to aid in the growth and survival of this infectious pathogen[1]. Fourteen of these effectors, including SidJ, have previously been shown to directly suppress the activities of other *Legionella* effectors[4,6,7]. SidJ opposes the toxicity of the SidE class of ubiquitin ligases (comprising SdeA, SdeB, SdeC and SidE) in yeast and mammalian cells. Deletion of SidJ in *Legionella* elicits substantial growth defects, probably owing to the unhinged toxicity of SidE enzymes[4]. In a recent report, SidJ has been shown to act as a deubiquitinase for canonical ubiquitination and for phosphoribosyl-linked ubiquitination mediated by SidE, which might account for the anti-SidE activity of SidJ[8]. However, we could not detect any intrinsic deubiquitinase activity of SidJ expressed in *Escherichia coli* or in mammalian cells (Extended Data Fig. 1a–c). Instead, we identified SidJ-associated proteins with deubiquitinase activities that co-precipitated with SidJ isolated from *Legionella* lysate, which potentially explains the previous reported observations regarding its anti-SidE activity (Extended Data Fig. 1c). SidE enzymes contain a mono-ADP-ribosyl transferase (mART) domain and a phosphodiesterase (PDE) domain, which sequentially perform ADP ribosylation of ubiquitin and substrate phosphoribosyl-linked ubiquitination, respectively[2,3]. In HEK293T cells, expression of SidJ with SdeA abolished the SdeA-mediated phosphoribosylation of ubiquitin, and nullified the ADP ribosylation of ubiquitin that is catalysed by the PDE-defective mutant SdeA(H277A) (which retains mART activity)[3] (Fig. 1a). Consistently, the yeast toxicity exerted by wild-type SdeA and the SdeA(H277A) mutant was rescued by co-expressing SidJ (Fig. 1b). Thus, we hypothesized that SidJ inhibits the first step of SdeA-mediated ubiquitination of substrates (that is, mono-ADP ribosylation of ubiquitin). As expected, although SdeA alone showed robust ε-NAD^+ hydrolysis, SdeA co-expressed with SidJ (hereafter, SidJ-treated SdeA) did not show mART activity (Fig. 1c). Lysate of HEK293T cells that express SidJ—but not lysate of untransfected cells or SidJ-expressing cell lysate that is depleted of SidJ—were able to block in vitro ε-NAD^+ hydrolysis mediated by SdeA (Fig. 1d). The addition of 10 mM ethylenediaminetetraacetic acid (EDTA) to lysates of SidJ-expressing HEK293T cells decreased the effect of these lysates on SdeA activity. Moreover, SidJ purified from *E. coli* was

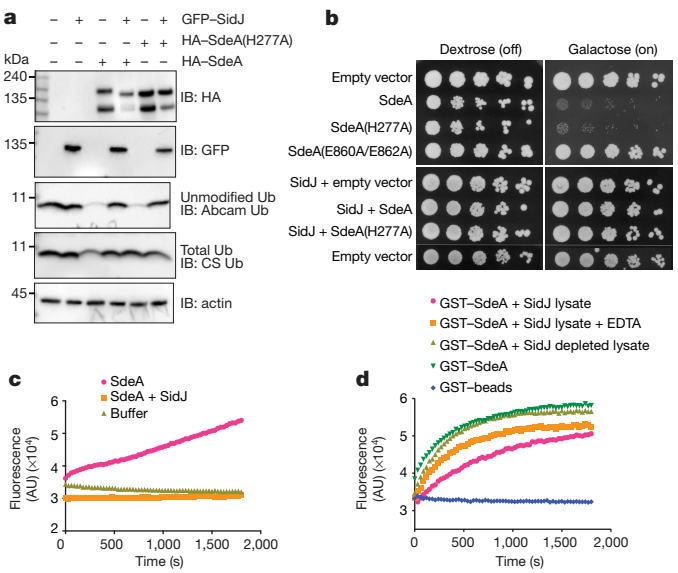


Fig. 1 | SidJ inhibits the ubiquitin–ADP ribosylation activity of SdeA. **a**, SidJ and SdeA constructs were expressed as indicated in HEK293T cells and ubiquitin modification was probed using the ubiquitin (Ub) antibodies Abcam Ub and Cell Signaling (CS) Ub, as previously described[3]. IB, immunoblot. **b**, Yeast strain W303 was transformed using the indicated combination of constructs (H277A, PDE defective; E860A/E862A, mART defective). Serial dilutions of transformed yeast were spotted on plates containing dextrose (repressing) or galactose (inducing). **c**, Purified SdeA, from HEK293T cells that express SdeA alone or in combination with SidJ, was used in ε-NAD^+ hydrolysis assays. The increase in the fluorescence indicates ubiquitin–ADP ribosylation[17]. **d**, Glutathione-*S*-transferase-tagged SdeA (GST–SdeA) purified from *E. coli* was incubated with HEK293T cell lysate that contained SidJ or was depleted of SidJ. SdeA was subsequently purified using glutathione agarose beads and used in ε-NAD^+ hydrolysis assays. Experiments in **a**–**d** were repeated three times independently with similar results. For gel source data, see Supplementary Fig. 1.

[1]Institute of Biochemistry II, Faculty of Medicine, Goethe University, Frankfurt am Main, Germany. [2]Buchmann Institute for Molecular Life Sciences, Goethe University, Frankfurt am Main, Germany. [3]European Molecular Biology Laboratory, Grenoble, France. [4]Department of Biology and Microbiology, South Dakota State University, Brookings, SD, USA. [5]Max Planck Institute of Biophysics, Frankfurt am Main, Germany. *e-mail: bhogaraju@embl.fr; dikic@biochem2.uni-frankfurt.de

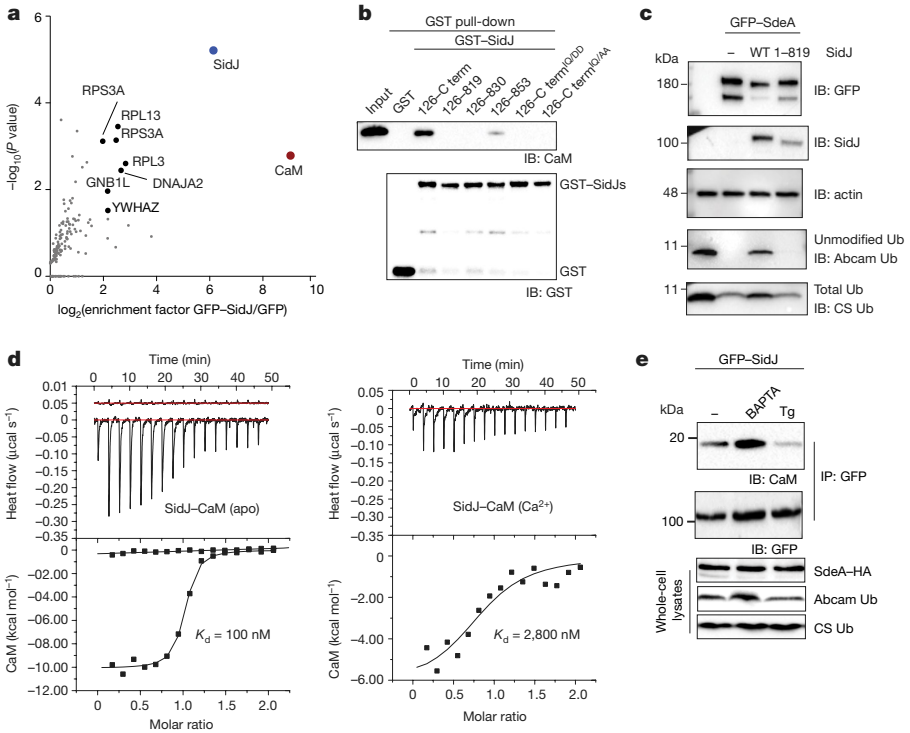


Fig. 2 | CaM is a host-specific factor that activates SidJ. a, GFP–SidJ was expressed in HEK293T cells. After immunoprecipitation of SidJ, the sample was analysed using mass spectrometry and enriched proteins were quantified using the MaxQuant label-free algorithm. $n = 3$ biologically independent experiments. Significant differences between samples were detected by a corrected, one-sided Student's *t*-test with a permutation-based false-discovery rate of 0.05. **b**, Various SidJ constructs were used to pull-down CaM to test the effect of mutations and deletions of the IQ motif. SidJ constructs that span residues 126 to 873, 126 to 819, 126 to 830 and 126 to 853 are labelled as 126–C term, 126–819, 126–830 and 126–853, respectively. The constructs labelled 126–C term$^{IQ/DD}$ and 126–C term$^{IQ/AA}$ are based on the SidJ construct 126–C-term, with residues I841 and Q842 mutated to aspartates or alanines, respectively. **c**, SdeA was expressed in HEK293T cells alone and in combination with wild-type (WT) SidJ or SidJ that lacks the IQ-motif region. Ubiquitin modification was followed using Abcam Ub and CS Ub antibodies. **d**, Isothermal titration calorimetry was performed to measure the affinity between SidJ and apo or Ca^{2+}-bound CaM. **e**, Interaction between GFP–SidJ and CaM was analysed using co-immunoprecipitation (IP) in low (treatment with BAPTA) or high (treatment with thapsigargin, Tg) cytosolic Ca^{2+} levels. Experiments in **b**–**e** were repeated three times independently with similar results. For gel source data, see Supplementary Fig. 1.

unable to catalyse such activities, which indicates that SidJ may be a metal-ion-dependent enzyme that requires one or more mammalian co-factors for its activity.

To identify the putative mammalian co-factor or co-factors that are necessary for the activity of SidJ, we expressed SidJ tagged with green fluorescent protein (GFP–SidJ) in HEK293T cells, and performed immunoprecipitation followed by mass spectrometry. Calmodulin (CaM) emerged as the strongest interactor of SidJ (Fig. 2a, Extended Data Fig. 2a). HHpred analysis of the SidJ sequence revealed a C-terminally located IQ motif, a well-known module that interacts with CaM[9], and its mutation (SidJ(I841A/Q842A) and SidJ(I841D/I841D)) or deletion (SidJ(ΔIQ)) resulted in a loss of binding to CaM (Fig. 2b). Co-expression of SidJ(ΔIQ) with SdeA did not lead to inhibition of SdeA activity in cells (Fig. 2c), nor of ε-NAD^+ hydrolysis catalysed by SdeA in vitro (Extended Data Fig. 2b). apo-CaM interacted with SidJ with a dissociation constant (K_d) of about 100 nM; by contrast, Ca^{2+}-loaded CaM bound to SidJ with a K_d of 2,800 nM, which indicates that Ca^{2+} binding to CaM reduces the strength of interaction between CaM and SidJ by nearly 30-fold (Fig. 2d). Treating HEK293T cells with the Ca^{2+} chelator BAPTA, or with the sarco- and endoplasmic-reticulum Ca^{2+}-ATPase pump inhibitor thapsigargin, demonstrated that the reduction of cytosolic free Ca^{2+} in cells increased the binding between SidJ and CaM, and correspondingly increased the inhibitory activity of SidJ towards SdeA (and vice versa) (Fig. 2e). Although global Ca^{2+} concentrations during *Legionella* infection do not considerably differ from those of uninfected cells (Extended Data Fig. 2c), we observed local dynamics of Ca^{2+} levels at the endoplasmic reticulum and at contact sites between the endoplasmic reticulum and *Legionella* (Extended Data Fig. 2d, Supplementary Video 1) that could have a role in the regulation of SidJ–CaM activity in vivo.

To unravel the mechanism by which SidJ affects SdeA activity, we tested whether SidJ adds an inactivating post-translational modification to SdeA. To this end, we immunoprecipitated SdeA and SidJ-treated SdeA from HEK293T cells, followed by quantitative mass spectrometry analysis. Although data analysis using MaxQuant and post-translational modification (PTM) discovery did not directly reveal any modification of SdeA, we noted that one of the SdeA tryptic peptides (residues 855–877)—which lines the active site of the mART domain—was severely underrepresented or not identified in the SidJ-treated SdeA samples, which suggests that this region of SdeA may undergo an uncommon modification by SidJ (Fig. 3a). Most of the region that spans between residues 855 and 877 is not accessible to solvent (Extended Data Fig. 2e)—except for residues 855–860, which form a solvent-accessible loop that contains the catalytic E860. To generate a peptide of this loop that was slightly longer but had a higher charge than the corresponding tryptic peptide, we used Lys-C to digest SidJ-modified SdeA and performed a PTM discovery analysis on the resulting peptides[10]. This revealed mono- and, to a lesser extent, di-glutamate conjugation to the E860 of SdeA (Fig. 3b, Extended Data Fig. 3a). Moreover, tandem mass tag (TMT) quantification of SdeA alone and SidJ-treated SdeA showed near to complete glutamylation of SdeA residue E860 in the latter condition (Fig. 3c). To recapitulate the glutamylation of SdeA in vitro, we treated SdeA with SidJ purified from *E. coli*, with or without CaM, in the presence of ATP, Mg^{2+} and L-glutamate. SidJ attenuated the ε-NAD^+ hydrolysis activity of SdeA in an ATP- and CaM-dependent manner (Extended Data Fig. 3b). Mass spectrometry analysis of in vitro SidJ reactions confirmed that the E860 of SdeA undergoes

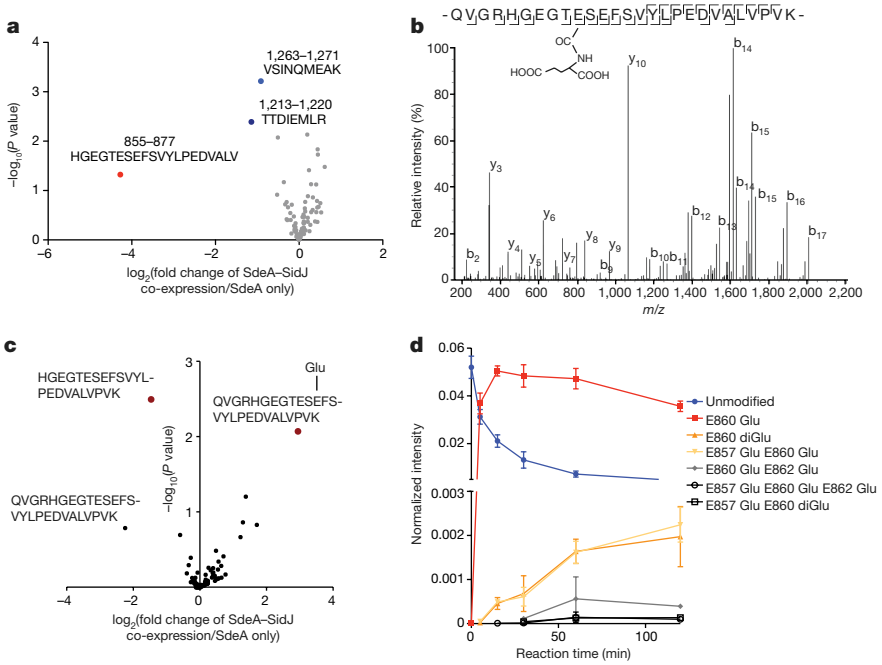


Fig. 3 | SidJ is a CaM-dependent glutamylase. a, Co-expression of SdeA with SidJ, and of SdeA alone, was performed in $n = 3$ biologically independent experiments. Samples were labelled with TMT six-plex reagent and analysed in one liquid chromatography–mass spectrometry run. Significant differences between samples were detected by a two-sided Student's *t*-test. The quantitative analysis of SdeA peptides under these conditions is represented in a volcano plot. **b**, Annotated mass spectra for glutamylation of E860 of SdeA. **c**, To obtain quantitative information about the modification, we purified GFP–SdeA expressed alone or co-expressed GFP–SdeA with SidJ, and digested with LysC. TMT quantification revealed close to quantitative conversion of the peptide that spanned the catalytic loop to its glutamylated (Glu) form. $n = 3$ biologically independent experiments. Significant differences between samples were detected by a two-sided Student's *t*-test. **d**, Intensities of different modified versions of the QVGRHGEGTESEFSVYLPEDVALVPVK peptide (the catalytic centre of the mART domain) are plotted as fraction of total SdeA intensity over the time of an in vitro reaction, which shows that E860 mono-glutamylation is the primary reaction of SidJ. In vitro glutamylation and label-free liquid chromatography–mass spectrometry analysis was performed in $n = 3$ biologically independent experiments. Data points are mean-centred; error bars indicate s.d.

glutamylation by SidJ–CaM in an ATP-dependent manner (Extended Data Fig. 4a–c). Mass spectrometry analysis of time-resolved in vitro glutamylation reactions revealed that SidJ–CaM targets SdeA preferentially by mono-glutamylation (Fig. 3d, Extended Data Fig. 5a–c). In prolonged reactions, we observed multiple modification states of SdeA, including glutamylation of E857 and E862. We did not detect polyglutamylation (defined as more than three glutamates) of SdeA by SidJ in vitro or in cells (Fig. 3d, Extended Data Figs. 4, 5), which indicates that SidJ–CaM is primarily a mono-glutamylating enzyme.

To gain insights into the mechanism of glutamylation by SidJ, we determined the structure of SidJ in complex with human apo-CaM (Extended Data Fig. 6a). Using single-particle cryo-electron microscopy (cryo-EM), we obtained a 3D reconstruction of the SidJ–CaM complex at a nominal resolution of 4.1 Å (Extended Data Figs. 6b–d, 7). Although we could build a partial de novo model using the electron-microscopy map, we used the high-resolution crystal structure of SidJ–CaM complex—which was deposited while our cryo-EM work was underway—as our initial model[11]. The crystal structure fit readily into the electron-microscopy map (Extended Data Fig. 6e) but there were some noticeable differences between the two, especially in the C-terminal domain of CaM (Extended Data Fig. 6f). We refined our model against the cryo-EM map, which considerably improved its fit into the density (Extended Data Fig. 8a, b, Supplementary Table 1). Local resolution analysis and visual inspection of the map revealed that most regions of SidJ and the N-terminal domain of CaM are well-resolved in our electron-microscopy map, with interpretable side-chain density visible for most amino acids in high-resolution regions (Extended Data Fig. 8b–d). Density for the C-terminal domain of CaM is less well-resolved, but the secondary structure elements could be positioned (Extended Data Fig. 8c). The structure of SidJ in complex with human CaM revealed the kinase-like catalytic domain of SidJ, an N-terminal α-helical domain and a distinct four-helix bundle that contains the IQ motif at the C terminus of SidJ, which mediates most of the interactions with CaM (Fig. 4a). SidJ and CaM buried a surface area of about 1,900 Å^2, in which the C-terminal domain of CaM semi-encircles the C-terminal helix of SidJ and the N-terminal domain of CaM makes extensive contacts with the N terminus of SidJ. Structurally, CaM binding to SidJ may stabilize the position of the N-lobe of the kinase-like domain, and thereby lead to the formation of a stable catalytic pocket. Our model is consistent with the crystal structure, including the overall architecture of the SidJ–CaM complex[11]. There is a noticeable movement of helices in the C-terminal domain of CaM compared to the crystal structure, probably because of the inherent differences between the structures of yeast and human CaM[12,13] (Fig. 4a). Notably, all the CaM structures in the RCSB Protein Data Bank (PDB) that are similar to CaM bound to SidJ are in apo conformation, which provides strong support for the notion that apo-CaM is the preferred conformation for SidJ binding. The crystal structure contains SidJ in complex with yeast Ca^{2+}-bound CaM; yeast CaM is approximately 60% similar to human CaM, and only has three Ca^{2+} binding sites (instead of four in human CaM)[12,13]. Given this, and our observation that apo-CaM is a better binder and activator of SidJ than the Ca^{2+}-bound CaM (Fig. 2d, e, Extended Data Fig. 3b), we propose that the architecture of SidJ–CaM observed in our cryo-EM structure represents the active SidJ–CaM glutamylase (see Extended Data Fig. 9a–c and Supplementary Discussion for more information).

Previous studies have shown that the ubiquitin ligase activity of the SidE family is attenuated one hour after infection with wild-type *L. pneumophila*, whereas the infection with the *ΔsidJ* strain of *Legionella* leads to prolonged activity of these ligases on the *Legionella*-containing vacuole (LCV)[4,14]. Calnexin-coated LCVs stained positive for glutamylation in A549 cells that were infected with wild-type *L. pneumophila* for 3 h, but not in cells that were infected with the *ΔsidJ* strain (Fig. 4b). We confirmed biochemically that SdeA is modified by

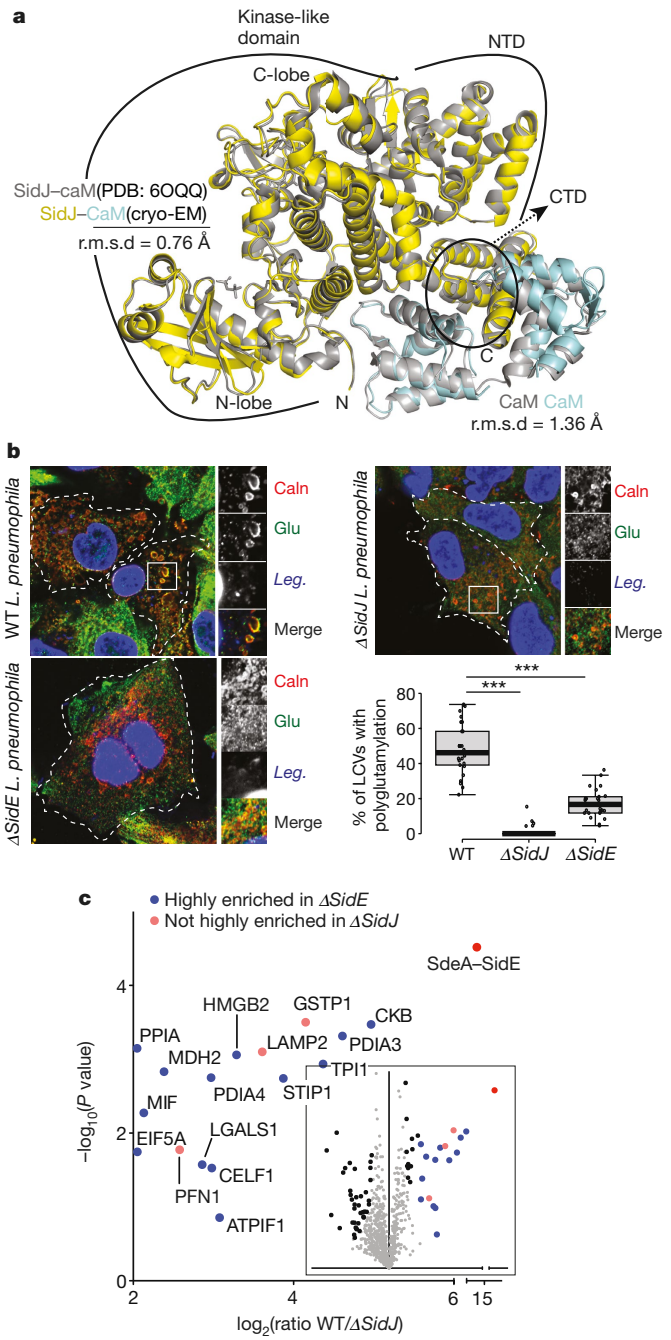


Fig. 4 | Glutamylation of SidE enzymes and host proteins during infection with *Legionella*. **a**, Comparison of the crystal structure of SidJ–yeast CaM (PDB: 6OQQ) with the cryo-EM structure showing root mean square deviation (r.m.s.d.) values of different regions. C, C terminus of SidJ; N, N terminus of SidJ; C-lobe, C-terminal lobe of the kinase-like domain; CTD, four-helix bundle that contains the IQ motif at the C terminus of SidJ; N-lobe, N-terminal lobe of the kinase-like domain; NTD, N-terminal α-helical domain of SidJ. **b**, A549 cells were infected with different strains of *L. pneumophila* for 3 h. Cells were fixed and immunostained with antibodies against calnexin (caln) and polyglutamylation (glu). DAPI staining marks the nucleus and cytosolic bacteria (*Leg.*). Number of LCVs (marked by calnexin) that are positive for polyglutamylation are counted in FIJI, and the percentage of polyglutamylated LCVs is plotted for cells infected with different strains of *Legionella*. Data represent 100 LCVs taken from 30 cells over $n = 3$ biologically independent experiments. Error bars indicate s.d. $***P < 0.001$, two-tailed, type-3 Student's *t*-test. *P* value $= 8.45 \times 10^{-15}$ (wild type versus *ΔsidJ*); *P* value $= 5.14 \times 10^{-11}$ (wild type versus *ΔsidE*). **c**, Glutamylated proteins were isolated from wild-type and *ΔsidJ Legionella* infection experiments using GT335 antibody, and quantified using mass spectrometry. Data are represented in volcano plot (inset) showing the most-enriched proteins in wild type. Grey circles, not significantly different; black, red and blue circles, significantly different. $n = 3$ biologically independent experiments. Significant differences between samples were detected by a corrected, two-sided Student's *t*-test with a permutation-based false-discovery rate of 0.05. Proteins with $\log_2$ ratio above two (mean) were labelled as highly enriched in wild-type compared to *ΔsidJ*-infected cells.

glutamylation during *Legionella* infection of mouse macrophages, in a SidJ-dependent manner (Extended Data Fig. 10a). SidJ-dependent glutamylation seen on LCVs was significantly lower in the *ΔsidE* strain (which lacks all four SidE ligases), as compared to cells infected with wild-type *L. pneumophila* (Extended Data Fig. 10b). Glutamylation on LCVs was not completely abolished in samples infected with the *ΔsidE* strain, which indicates that SidJ may target additional proteins for glutamylation during *Legionella* infection. To explore this, we immunoprecipitated glutamylated proteins from cells infected with wild-type or *ΔsidJ Legionella* and performed quantitative mass-spectrometry analysis (Fig. 4c). SdeA and SidE are the most enriched protein group, which demonstrates that the approach we adopted is suitable for finding bona fide substrates of SidJ glutamylation. Apart from SdeA, we observed several host proteins that are significantly enriched. To rule out the possibility that these proteins could have simply been co-immunoprecipitated with glutamylated SdeA, we did a similar quantitative analysis between cells infected with *ΔsidE* and *ΔsidJ* strains of *L. pneumophila* and found that—with the exception of three proteins (LAMP2, GSTP1 and LGALS1)—all of the potential targets of SidJ are also significantly enriched in the quantification of infection with *ΔsidE* versus *ΔsidJ L. pneumophila* (Extended Data Fig. 10c). These data show that SidJ has additional glutamylation targets, which may explain why the deletion of SidJ leads to an intracellular growth defect that is more severe than that following the deletion of all SidE enzymes[4,5,14].

In conclusion, SidJ is a CaM-dependent glutamylase that antagonizes the ubiquitin ligase activity of SidE enzymes, and their associated cellular toxicity, during *Legionella* infection. Despite the co-existence of SidJ and SidE enzymes in *Legionella*[4,5,14], SidJ glutamylase activity is spatially regulated and triggered only in the host cells by interacting with the eukaryote-specific CaM protein. The toxins oedema factor (of *Bacillus anthracis*) and CyaA (of *Bordetella pertussis*) also use CaM as a co-factor to exert toxic adenylate cyclase activities in host cells[15,16]. This identifies CaM as having a potentially important role in bacterial infection. The observed effects of levels of intracellular Ca^{2+} on the interaction between SidJ and CaM, and on the glutamylase activity of SidJ, present an additional level of regulation that might be instrumental in the spatiotemporal modulation of the ubiquitin ligase activity of SidE enzymes. The finding that SidJ can also mediate the glutamylation of several host proteins (in addition to bacterial effectors) offers a molecular explanation for the observed broader role of SidJ, as compared to SidE enzymes, in *Legionella* proliferation in host amoebae and in macrophages[4,5,14]. Future studies into these targets may shed light on as-yet unexplored host–pathogen interactions in *Legionella* infection.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1440-8.

Received: 25 April 2019; Accepted: 9 July 2019;
Published online 22 July 2019.

1. Hubber, A. & Roy, C. R. Modulation of host cell function by *Legionella pneumophila* type IV effectors. *Annu. Rev. Cell Dev. Biol.* **26**, 261–283 (2010).
2. Qiu, J. et al. Ubiquitination independent of E1 and E2 enzymes by bacterial effectors. *Nature* **533**, 120–124 (2016).

3. Bhogaraju, S. et al. Phosphoribosylation of ubiquitin promotes serine ubiquitination and impairs conventional ubiquitination. *Cell* **167**, 1636–1649. e13 (2016).
4. Jeong, K. C., Sexton, J. A. & Vogel, J. P. Spatiotemporal regulation of a *Legionella pneumophila* T4SS substrate by the metaeffector SidJ. *PLoS Pathog.* **11**, e1004695 (2015).
5. Liu, Y. & Luo, Z.-Q. The *Legionella pneumophila* effector SidJ is required for efficient recruitment of endoplasmic reticulum proteins to the bacterial phagosome. *Infect. Immun.* **75**, 592–603 (2007).
6. Havey, J. C. & Roy, C. R. Toxicity and SidJ-mediated suppression of toxicity require distinct regions in the SidE family of *Legionella pneumophila* effectors. *Infect. Immun.* **83**, 3506–3514 (2015).
7. Urbanus, M. L. et al. Diverse mechanisms of metaeffector activity in an intracellular bacterial pathogen, *Legionella pneumophila*. *Mol. Syst. Biol.* **12**, 893 (2016).
8. Qiu, J. et al. A unique deubiquitinase that deconjugates phosphoribosyl-linked protein ubiquitination. *Cell Res.* **27**, 865–881 (2017).
9. Bähler, M. & Rhoads, A. Calmodulin signaling via the IQ motif. *FEBS Lett.* **513**, 107–113 (2002).
10. Solntsev, S. K., Shortreed, M. R., Frey, B. L. & Smith, L. M. Enhanced global post-translational modification discovery with MetaMorpheus. *J. Proteome Res.* **17**, 1844–1851 (2018).
11. Black, M. H. et al. Bacterial pseudokinase catalyzes protein polyglutamylation to inhibit the SidE-family ubiquitin ligases. *Science* **364**, 787–792 (2019).
12. Starovasnik, M. A., Davis, T. N. & Klevit, R. E. Similarities and differences between yeast and vertebrate calmodulin: an examination of the calcium-binding and structural properties of calmodulin from the yeast *Saccharomyces cerevisiae*. *Biochemistry* **32**, 3261–3270 (1993).
13. Ishida, M. et al. The solution structure of apocalmodulin from *Saccharomyces cerevisiae* implies a mechanism for its unique Ca^{2+} binding property. *Biochemistry* **41**, 15536–15542 (2002).
14. Bardill, J. P., Miller, J. L. & Vogel, J. P. IcmS-dependent translocation of SdeA into macrophages by the *Legionella pneumophila* type IV secretion system. *Mol. Microbiol.* **56**, 90–103 (2005).
15. Drum, C. L. et al. Structural basis for the activation of anthrax adenylyl cyclase exotoxin by calmodulin. *Nature* **415**, 396–402 (2002).
16. Wolff, J., Cook, G. H., Goldhammer, A. R. & Berkowitz, S. A. Calmodulin activates prokaryotic adenylate cyclase. *Proc. Natl Acad. Sci. USA* **77**, 3841–3844 (1980).
17. Kalayil, S. et al. Insights into catalysis and function of phosphoribosyl-linked serine ubiquitination. *Nature* **557**, 734–738 (2018).

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized and investigators were not blinded to allocation during experiments and outcome assessment.

Protein purification. GST-tagged protein constructs of SidJ and SdeA were transformed into BL21(DE3) and grown in LB medium until the optical density at 600 nm (OD_{600}) reached 0.6. Cultures were induced using 0.5 mM isopropyl D-thiogalactopyranoside (IPTG) and allowed to grow overnight at 18 °C. T7-expressing *E. coli* cells were transformed with pGEX-6P-1-SidJ constructs and grown in LB medium containing ampicillin. Cells were grown until the OD_{600} reached 0.6–0.8 at 37 °C, induced with 0.2 mM IPTG and further grown at 18 °C. Collected cells were resuspended in lysis buffer (50 mM Tris-HCl pH 7.5 and 150 mM NaCl), sonicated and centrifuged at 13,000 rpm. The supernatant was incubated for 2 h at room temperature with glutathione-*S*-sepharose pre-equilibrated with washing buffer (50 mM Tris-HCl pH 7.5, 500 mM NaCl). Non-specific proteins were washed with washing buffer and GST–SidJ was eluted with elution buffer (50 mM Tris-HCl pH 8.0, 50 mM NaCl, 15 mM reduced glutathione). Proteins were buffer-exchanged back into the lysis buffer and stored. N-terminally His-tagged CaM was purified with Ni-NTA agarose and the following buffers: lysis buffer (50 mM Tris-HCl pH 7.5 and 150 mM NaCl, 10 mM imidazole), wash buffer (50 mM Tris-HCl pH 7.5 and 500 mM NaCl, 10 mM imidazole) and elution buffer (50 mM Tris-HCl pH 7.5 and 150 mM NaCl, 300 mM imidazole). Eluted CaM was further purified with size-exclusion chromatography (Superdex 75 10/300 GL, GE Healthcare) pre-equilibrated with Ca^{2+}-containing storage buffer (50 mM Tris-HCl pH 7.5 and 150 mM NaCl, 2 mM $CaCl_2$). To prepare apo-CaM, purified CaM was incubated with 10 mM EGTA for 2 h and buffer-exchanged with size-exclusion chromatography (Superdex 75 10/300 GL, GE Healthcare) pre-equilibrated with storage buffer (50 mM Tris-HCl pH 7.5 and 150 mM NaCl). Linkage-specific di-ubiquitin chains were purchased from UbiQ.

ε-NAD^+ hydrolysis assays. SdeA-containing samples are from HEK293T cells that transiently express either GFP–SdeA alone or in combination with haemagglutinin-tagged SidJ (HA–SidJ) or HA–SidJ(ΔIQ). SdeA-containing samples of in vitro glutamylation assays were also used for ε-NAD^+ hydrolysis assays. Reaction mixture contained 1 mM ε-NAD^+, 50 μg of ubiquitin and SdeA-containing samples were diluted to 100 μl in a buffer containing 50 mM Tris pH 7.5, 50 mM NaCl. Reaction was initiated by adding ε-NAD^+ and the hydrolysis was followed using a plate reader operating under fluorescence kinetic measurement mode, with excitation wavelength 300 nm and emission wavelength 410 nm. Measurements were taken of each sample every 30 s for the indicated time points.

GST-pulldown assays. GST or GST–SidJ were incubated with 30 μl of glutathione-*S*-sepharose beads (GE Healthcare) pre-equilibrated with binding buffer (50 mM Tris-HCl pH 7.5 and 150 mM NaCl) for 30 min. apo-CaM or Ca^{2+}-bound CaM were added to this and further incubated. Unbound proteins were washed with washing buffer (50 mM Tris-HCl pH 7.5 and 150 mM NaCl, 0.5% (v/v) NP-40), and samples were analysed by SDS–PAGE followed by immunoblotting. For the Ca^{2+}-bound CaM, all the buffers were supplemented with 2 mM $CaCl_2$.

Liquid chromatography–mass spectrometry analyses. For mass-spectrometric analysis, *Legionella*-purified SidJ was separated on a 1D gel, the gel was stained with Coomassie and cut into four fractions, which were subjected to in-gel digestion as previously described[18]. eGFP-tagged SdeA was immunopurified with anti-GFP beads (Chromotek), weak binding proteins were removed by 4 washes with 4 M urea buffer. Purified SdeA was denatured by boiling for 5 min in 2% sodium deoxycholate, 5 mM TCEP and 20 mM chloroacetamide in 50 mM Tris pH 8. Afterwards, SdeA was digested on beads either overnight with 0.5 μg trypsin after dilution of sodium deoxycholate to 1%, or for 30 min with LysC. The digestions were stopped by addition of 4 volumes 1% TFA in isopropanol and the peptides were desalted and enriched by SDB-RPS Stage-Tips. Proteins from in vitro glutamylation reactions were digested in solution with LysC and prepared as described for immunopurified samples. To obtain quantitative information, peptides were either labelled with TMT 6plex reagent (ThermoFisher) and further purified with C18 stagetips[19] or analysed label-free. Proteins from glutamylation immunoprecipitations were prepared in 2–4 M urea buffer as described for immunopurified samples, but without boiling the samples and with C18 Stage-Tip enrichment.

Peptides were separated on an in-house-designed C18 column (20-cm length, 75-μm inner diameter, 1.9-μm particle size) by an Easy n-LC 1200 (Thermo Fisher) and directly injected in a QExactive-HF or—in case of TMT samples—into a Fusion Lumos mass spectrometer (Thermo Fisher), and analysed in data-dependent mode.

Data analysis was done with MaxQuant 1.65[20]. In brief, samples were searched against the Uniprot *L. pneumophila* database (for SidJ purified from *Legionella*) or the human Swissprot database supplemented with the SdeA and SidJ sequences (for immunoprecipitation and in vitro samples) and eventually quantified by the TMT 6plex option. Glutamylated peptides were verified by manual interpretation of spectra. Quantitative changes were further analysed by unpaired *t*-tests in Perseus 1.65[21]. Global PTM discovery analysis was performed with MetaMorpheus[10]. For comparison of differentially modified peptide species, the precursor intensities of different charge states were summed up and normalized by the total intensity of the protein.

Deubiquitination assays. For conventional deubiquitinase assay, GST, USP2 and GST–SidJ were activated with activation buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM DTT), and incubated with linkage specific di-ubiquitin chains for 1 h at 37 °C, and samples were analysed by SDS–PAGE followed by immunoblotting. For the phosphoribosyl-linked ubiquitin deubiquitinase assay, phosphoribosyl-linked ubiquitinated Rtn4 peptides were incubated with GST or GST–SidJ for 1 h at 37 °C.

Immunocytochemistry and confocal imaging. For immunocytochemistry, cells were fixed with 4% paraformaldehyde followed by permeabilization in 10% fetal bovine serum, PBS and 0.1% saponin for 60 min, followed by overnight staining in primary antibody at 4 °C and 60 min incubation in secondary antibody at room temperature. Confocal imaging was done using the Zeiss LSM780 microscope system. Ar-ion laser (Alexa Fluor 488 with the 488-nm line), a He Ne laser (for Alexa Fluor 546 with the 543-nm line) and violet laser (for DAPI) were used with 63 × 1.4 NA oil-immersion objective. Image analyses were done using FIJI. To count calnexin-positive LCVs, (which are also positive for glutamylation), 25-μm^2 regions of interest (ROIs) at the perinuclear region of each cell were analysed. This was done for 30 cells taken from 3 biologically independent experiments. The Coloc2 plugin in FIJI was used in 50-μm^2 ROIs to calculate the Manders coefficient to quantify the colocalization between calnexin-marked LCVs and polyglutamylation. This was done for 80 ROIs from 20 cells. For live-cell imaging, cells were infected with *L. pneumophila* in carbon-dioxide-independent medium, maintaining the stage thermostat at 37 °C and with a 5% CO_2 supply. Images were recorded for 2 min, at 1-s intervals.

Calcium measurements using Fura2-AM. Cells were loaded with 2.5 μM Fura2AM in Tyrodes buffer (25 mM HEPES (pH 7.4), 5 mM potassium chloride, 140 mM sodium chloride, 2 mM magnesium chloride, 6 g/l glucose) for 30 min at 37 °C. The cells were washed and collected in PBS, followed by the measurement of fluorescence at 510 nm with the Tecan Infinite M200 Pro plate reader after alternate excitation at 340 nm and 380 nm (ref. [22]). Calcium levels were calculated using the formula $[Ca^{2+}] = K_d[(R - R_{min})/(R_{max} - R)] \times Sf$. R represents the ratio of fluorescence intensity at 340 nm and 380 nm[22]. R_{max} and R_{min} were calculated from cells treated with 1% digitonin (Ca^{2+} saturation) and 1% digitonin + 2 mM EGTA (absence of Ca^{2+}) respectively. Sf is the scaling factor (fluorescence intensity at 380-nm excitation in the absence of Ca^{2+} and at Ca^{2+} saturation).

***L. pneumophila* infection and preparation of lysates from infected cells.** *L. pneumophila* strains (*L. pneumophila* lp02) were grown for 3 days on *N*-(2-acetamido)-2-amino-ethanesulfonic acid (ACES)–buffered charcoal–yeast (BCYE) extract agar, at 37 °C. Bacteria were grown for 20 h in ACES medium before infection. Bacterial cultures of OD_{600} 3.2–3.6 were used to infect RAW 264.7 cells (multiplicity of infection (MOI) of 1:10) and A549 cells (MOI of 1:20).

After the infection, the cells were pelleted at 800*g* followed by treatment with 0.05% digitonin in KHEM buffer (10 mM HEPES (pH 7.2), 140 mM potassium chloride, protease inhibitor cocktail) for 10 min at room temperature, followed by centrifugation at 13,000*g* for 10 min. The supernatant contains purely cytosolic proteins (such as tubulin), and the pellet contains all cellular membranes (including LCVs). The pellet was then lysed in buffer containing 1% Triton X-100 and used as an input for immunoprecipitation with GT335 antibody.

Cell lines. HEK293T (ATCC CRL-3216) and A549 (ATCC CCL-185) cells were cultured in DMEM supplemented with 10% FBS, 100 IU/ml penicillin and 100 mg/ml streptomycin (penicillin–streptomycin) at 37 °C, 5% CO_2. Raw264.7 macrophages (ATCC TIB-71) were cultured in RPMI supplemented with 10% FBS. All cell lines were verified by general morphology or short tandem repeat analysis and found to contain no mycoplasma using a PCR test.

Western blotting and immunoprecipitation. Four to twenty per cent Tris glycine gradient gels (Biorad) were used for SDS–PAGE, followed by western blot. For immunoprecipitation, cells were lysed in immunoprecipitation buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1% Triton X-100, 1 mM PMSF, protease inhibitor cocktail (Sigma Aldrich)), mixed with 15 μl protein A/G agarose beads (SantaCruz Biotechnology) and 3 μg anti-polyglutamylation (GT335) antibody, and incubated for 4 h at 4 °C while being subjected to end-to-end rotation. The beads were washed twice in immunoprecipitation buffer containing 400 mM NaCl. Proteins were eluted by boiling with 2× gel loading dye, followed by western blot.

In vitro glutamylation assays. One microgram of SdeA and 2 μg of GST–SidJ were included in the reaction mixtures. Where mentioned, 2 μg of apo- or Ca^{2+}-bound CaM, as well as 0.5 mM ATP and 0.5 mM glutamic acid was added to the reactions. All reactions were performed in 50 mM Tris pH 9, 2.5 mM $MgCl_2$, 0.5 mM TCEP in a final volume of 20 μl. The reaction was initiated by the addition of ATP and incubated at 30 °C for 1 h. For time-course mass spectrometry

experiments checking the specificity of the reaction, the amount of glutamic acid and ATP was reduced to 0.05 mM, and the reaction was carried out for indicated time points.

Expression and purification of SidJ–CaM complex for cryo-EM. Recombinant overexpression of full-length SidJ and CaM were achieved through co-expression using *E. coli* OneShot BL21 Star (DE3) cells (Invitrogen). The cells were cultivated at 37 °C in lysogeny broth supplemented with 100 μg/ml ampicillin, 25 μg/ml kanamycin and 34 μg/ml chloramphenicol. Expression was induced at $OD_{600} = 0.6$ using a final concentration of 0.5 mM isopropyl-β-D-1-thio-galactopyranoside (IPTG), and the cultures were further left to grow at 18 °C. Cells were collected the next day by centrifugation and lysed by sonication in lysis buffer (50 mM Tris pH 7.5, 300 mM NaCl, 5% (v/v) glycerol), supplied with an EDTA-free protease inhibitor cocktail tablet (Roche Applied Science) and 20 μg/ml DNase 1. The lysate was clarified using centrifugation and filtered using a 5-μM filter membrane, before applying the lysate to TALON metal-affinity resin beads (ClonTech Laboratories) pre-equilibrated in lysis buffer. The lysate was left to incubate on the beads at 4 °C for 1 h, and the flow-through was removed after gentle centrifugation at 300*g* for 2 min. The beads were washed 3 times with lysis buffer and an incubation time of 10 min at 4 °C, and subsequent gentle centrifugation during each step. The elutions were performed using lysis buffer that was supplied with increasing imidazole concentrations of 10 mM, 50 mM, 100 mM and 200 mM, an incubation time of 10 min, and subsequent gentle centrifugation. The purest fraction as determined by SDS–PAGE was placed in Spectra/Por 1 RC Dialysis Membrane Tubing (Spectrum), and His-3C protease in a 1:50 molar ratio was added for tag cleavage. The sample was then dialysed in dialysis buffer (50 mM Tris pH 7.5, 150 mM NaCl, 5% glycerol, 5 mM EDTA) overnight and loaded onto a HiLoad Superdex S200 10/300 GL (GE Healthcare) column equilibrated in gel filtration buffer (10 mM HEPES pH 7.5, 50 mM NaCl, 0.5 mM TCEP) for size-exclusion chromatography, and the purest fractions (as determined by SDS–PAGE) were used for grid preparation.

Expression and purification of SidJ(99–C) and full-length CaM for isothermal titration calorimetry. Recombinant overexpression of the SidJ construct that spans residues 99 to 873 (99–C) and full-length CaM were achieved through expression using *E. coli* OneShot BL21 Star (DE3) cells (Invitrogen). The cells were cultivated at 37 °C in lysogeny broth supplemented with 25 μg/ml kanamycin and 34 μg/ml chloramphenicol for SidJ(99–C) and 100 μg/ml ampicillin and 34 μg/ml chloramphenicol for CaM. Both SidJ(99–C) and CaM were expressed, lysed and purified using the same protocol as the SidJ–CaM complex described in 'Expression and purification of SidJ–CaM complex for cryo-EM', but CaM was incubated in lysis buffer with 4mM EDTA, added at 4 °C for 2 h, before size-exclusion chromatography, and neither protein was cleaved or dialysed.

Isothermal titration calorimetry experiments. The SidJ(99–C) and full-length CaM purified as described in 'Expression and purification of SidJ(99–C) and full-length CaM for isothermal titration calorimetry' were concentrated to 20 μM and 200 μM, respectively, using Amicon Ultra centrifugal filters (Merck Millipore). The isothermal titration calorimetry experiments were performed using a MicroCal ITC200 (Malvern), with a sample volume of 350 μl and a ligand volume of 75 μl. All experiments were performed at 20 °C using an initial injection volume of 1 μl, and all subsequent injections with a volume of 2.5 μl, with 5-min injection intervals. For baseline measurements, the 200 μM CaM was titrated into the gel filtration buffer (10 mM HEPES pH 7.5, 50 mM NaCl, 0.5 mM TCEP). For the apo-CaM measurement, 200 μM CaM was titrated into the sample chamber containing 20 μM SidJ(99–C). For the $CaCl_2$-enriched CaM measurement, 50 mM $CaCl_2$ was added to both CaM and SidJ(99–C) to reach a final concentration of 3 mM $CaCl_2$, and left to incubate on ice for 2 h before the experiment.

Data baseline subtraction, analysis and the determination of dissociation constants were performed using Origin 7.0. The first injection was excluded from analysis in each experiment.

Electron microscopy. Cryo-EM grids were prepared using Vitrobot MK IV (Thermo Fisher) operated at 100% humidity at 4 °C. Two microlitres of sample was applied to each side of an UltrAufoil 300 mesh, 1.2/1.3 grid, glow-discharged using Pelco EasyGlow at 25 mA for 30 s, and blotted for 2 s before immediate plunge-freezing into liquid ethane. Samples were imaged using Glacios microscope (Thermo Fisher) operated at 200 kV with Falcon III direct electron detector. Two thousand four hundred and forty-one movies were acquired in counting mode (defocus range of −0.5 μm to -2.5 μm) with magnified pixel size of 0.96 Å at a dose rate 0.9 e $Å^{-2}s^{-1}$, and the total dose of 40 e $Å^{-2}$ fractionated into 60 movie frames.

Electron-microscopy data processing. Motion correction and contrast-transfer function (CTF) parameter estimation was performed in WARP[23] using 5 × 5 patches, followed by particle-picking with the BoxNet2Mask_21080918 model. Coordinates of a total of 1,500,000 particles were imported into Relion3[24] and extracted with a binning factor of 2. After 2 rounds of reference-free 2D classification, 742,000 particles were subjected to 3D classification with an initial model generated ab initio within Relion3. A subset of 369,000 particles from the best 3D class was re-extracted without binning, classified again, and a 108,000-particle subset was refined to 4.5-Å resolution. Further 3D classification without image alignment (with $T = 8$) allowed isolation of a 20,000-particle subset, which was refined to 4.1-Å resolution. Finally, the CTF refinement and beam-tilt correction was performed, followed by Bayesian particle polishing[25], but the quality of the reconstruction did not benefit substantially from these procedures (Extended Data Fig. 7).

Model building and refinement. The average resolution of the reconstruction based on gold-standard Fourier shell correlation (FSC) is 4.1 Å. However, many areas of the map show clear density for amino acid side chains (Extended Data Fig. 7d), which indicates that these parts of the map might be suitable for de novo model building. While we were attempting model building, a crystal structure of SidJ bound to yeast CaM was reported at 2.1-Å resolution (PDB 6OQQ). This structure was rigid-body-fitted into the cryo-EM density in Chimera and the yeast CaM was replaced with an apo form of human CaM (PDB 2IX7). The structure was subsequently refined against the cryo-EM map with Refmac5[26] implemented in CCP-EM[27] using secondary structure restraints from Prosmart[28]. Refinement and validation statistics are summarized in Supplementary Table 1.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Mass spectrometry data are available from the Proteomics Identification (PRIDE) database (https://www.ebi.ac.uk/pride/archive/) with the dataset identifier[29] PXD014362. Cryo-EM structure coordinates are available from the Protein Data Bank (PDB) and the Electron Microscopy Data Bank (EMDB) under accession codes 6S5T and EMD-10100, respectively. Full gel source data can be found in Supplementary Fig. 1. The data that support the findings of this study are available from the corresponding authors upon request.

18. Grumati, P. et al. Full length RTN3 regulates turnover of tubular endoplasmic reticulum via selective autophagy. *eLife* **6**, e25555 (2017).
19. Rappsilber, J., Ishihama, Y. & Mann, M. Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal. Chem.* **75**, 663–670 (2003).
20. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).
21. Tyanova, S. et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).
22. Grynkiewicz, G., Poenie, M. & Tsien, R. Y. A new generation of Ca^{2+} indicators with greatly improved fluorescence properties. *J. Biol. Chem.* **260**, 3440–3450 (1985).
23. Tegunov, D. & Cramer, P. Real-time cryo-EM data pre-processing with Warp. Preprint at https://www.biorxiv.org/content/10.1101/338558v1 (2018).
24. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).
25. Zivanov, J., Nakane, T. & Scheres, S. H. W. A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. *IUCrJ* **6**, 5–17 (2019).
26. Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D* **67**, 355–367 (2011).
27. Burnley, T., Palmer, C. M. & Winn, M. Recent developments in the CCP-EM software suite. *Acta Crystallogr. D* **73**, 469–477 (2017).
28. Nicholls, R. A., Fischer, M., McNicholas, S. & Murshudov, G. N. Conformation-independent structural comparison of macromolecules with ProSMART. *Acta Crystallogr. D* **70**, 2487–2499 (2014).
29. Perez-Riverol, Y. et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* **47**, D442–D450 (2019).
30. Scheres, S. H. W. & Chen, S. Prevention of overfitting in cryo-EM structure determination. *Nat. Methods* **9**, 853–854 (2012).
31. Brown, A. et al. Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. *Acta Crystallogr. D* **71**, 136–153 (2015).

Acknowledgements We thank J. Vogel for the *ΔsidJ Legionella* strain; Z.-Q. Luo for SidJ proteins purified from *Legionella*, and for providing *ΔsidE* and wild-type *Legionella* strains; S. Rodriguez and S. Gharbi for technical assistance; Y. Liu for help with deubiquitination assays; D. Höller, H. Marei and K. Koch for critical comments on the manuscript; and S. Knapp and V. Doetsch for discussion and advice. This work was supported by the DFG-funded Collaborative Research Centre on Selective Autophagy (SFB 1177), by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 742720), by the DFG-funded Cluster of Excellence 'Macromolecular Complexes' (EXC115) and by the DFG-funded SPP 1580 program 'Intracellular Compartments as Places of Pathogen-Host-Interactions' (to I.D.).

Author contributions S.B. and I.D. conceived and supervised the project. S.B. performed mammalian SidJ and SdeA co-expression and ε-NAD^+ hydrolysis

assays. F.B. performed all the mass spectrometry experiments with the help of V.M. and S.B. in sample preparation. R.M. performed *Legionella* infection experiments and prepared samples for mass spectrometry. M.A. performed in vitro ε-NAD^+ hydrolysis assays, isothermal titration calorimetry and protein purifications for biochemistry and cryo-EM. M.M.P. and W.P.G. prepared cryo-EM grids, collected and processed data and performed the model refinement. J.L.-M. performed yeast experiments. S.K. performed protein purifications. D.S. performed protein purifications and in vitro SidJ–CaM pulldown experiments. S.B. and I.D. analysed the data and wrote the manuscript with input from all the co-authors.

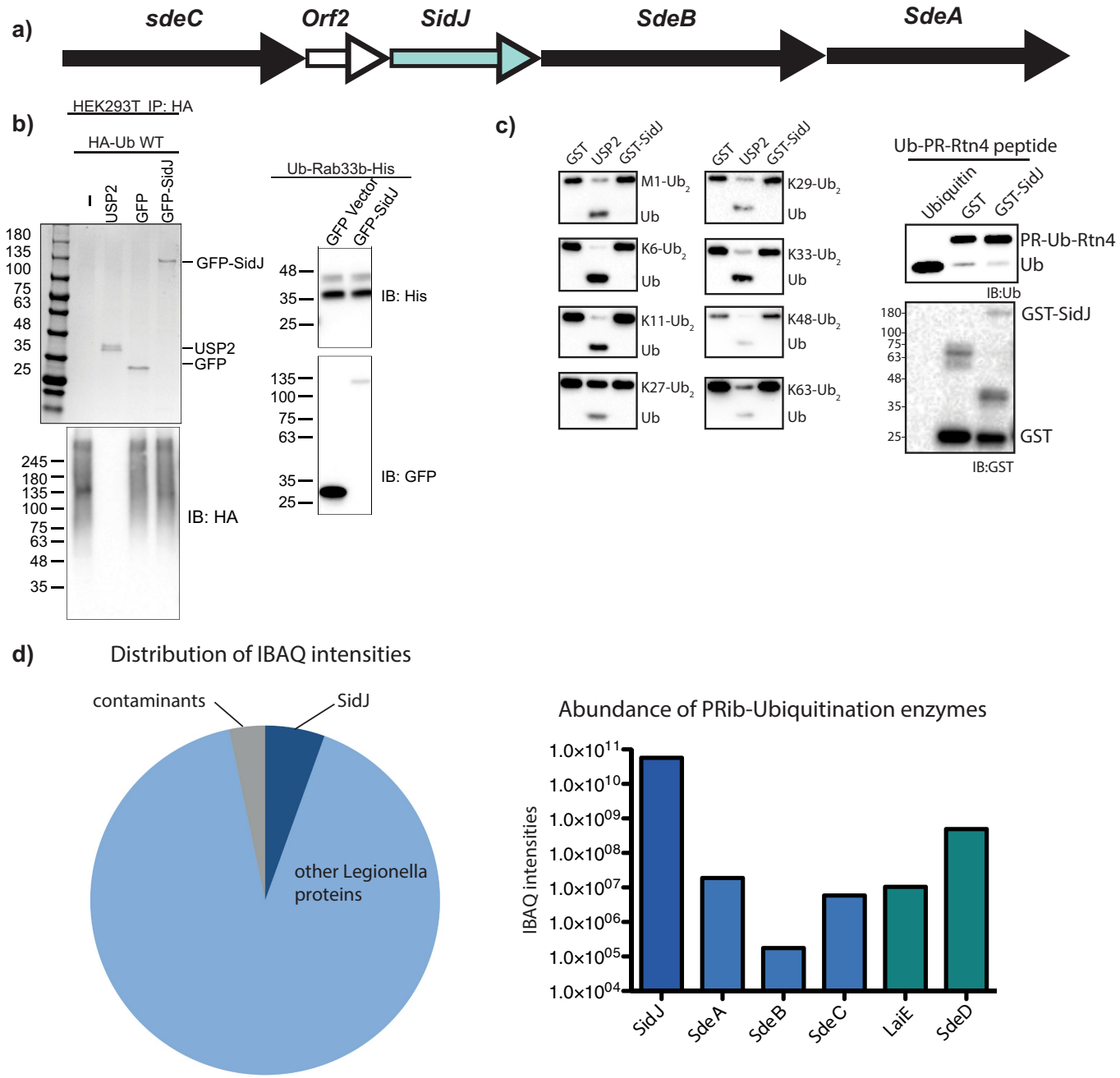
Competing interests The authors declare no competing interests.

Additional information

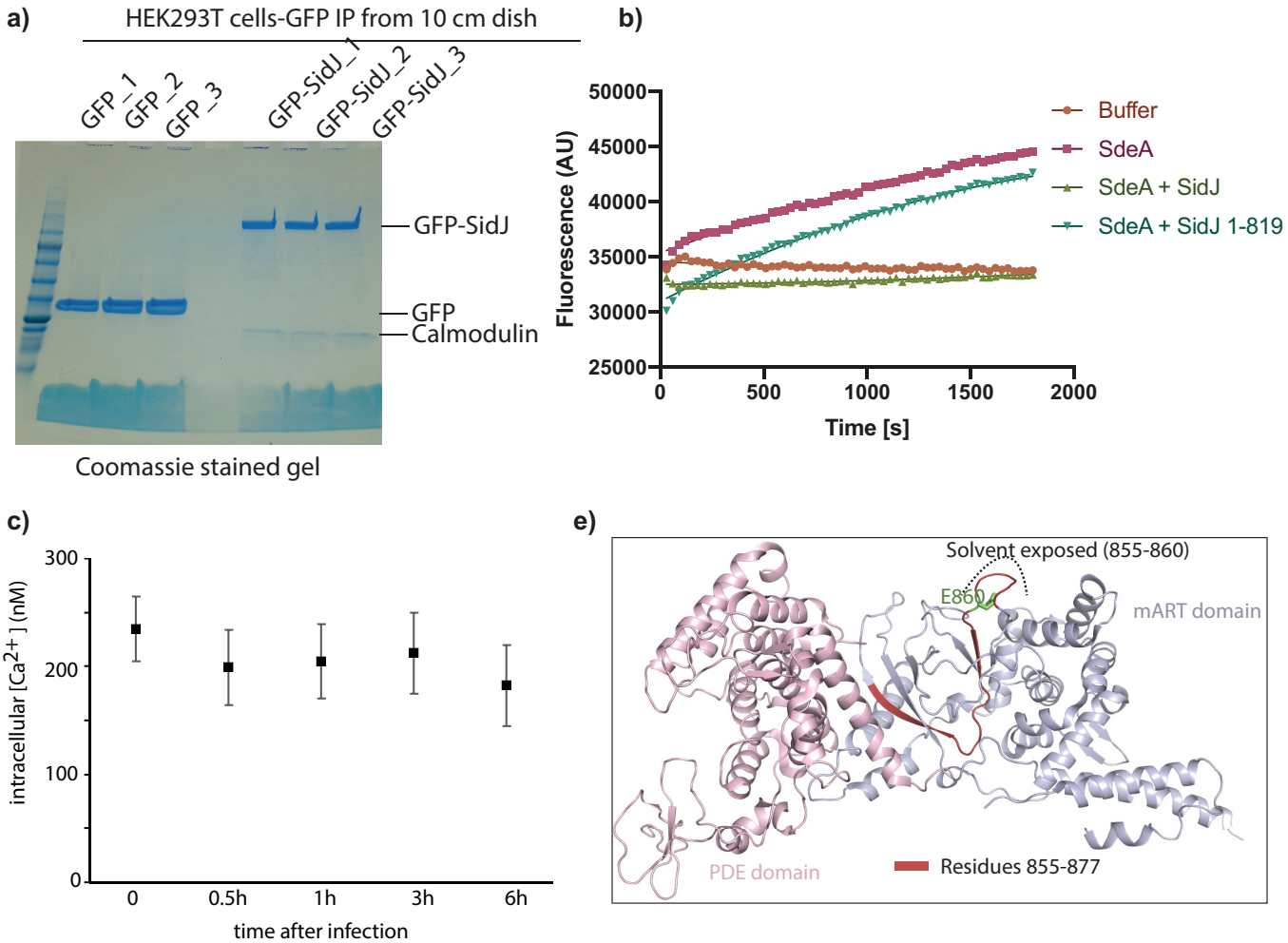
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1440-8.

Correspondence and requests for materials should be addressed to S.B. or I.D.

Peer review information *Nature* thanks Friedrich Förster, Elizabeth Hartland, Carsten Janke and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

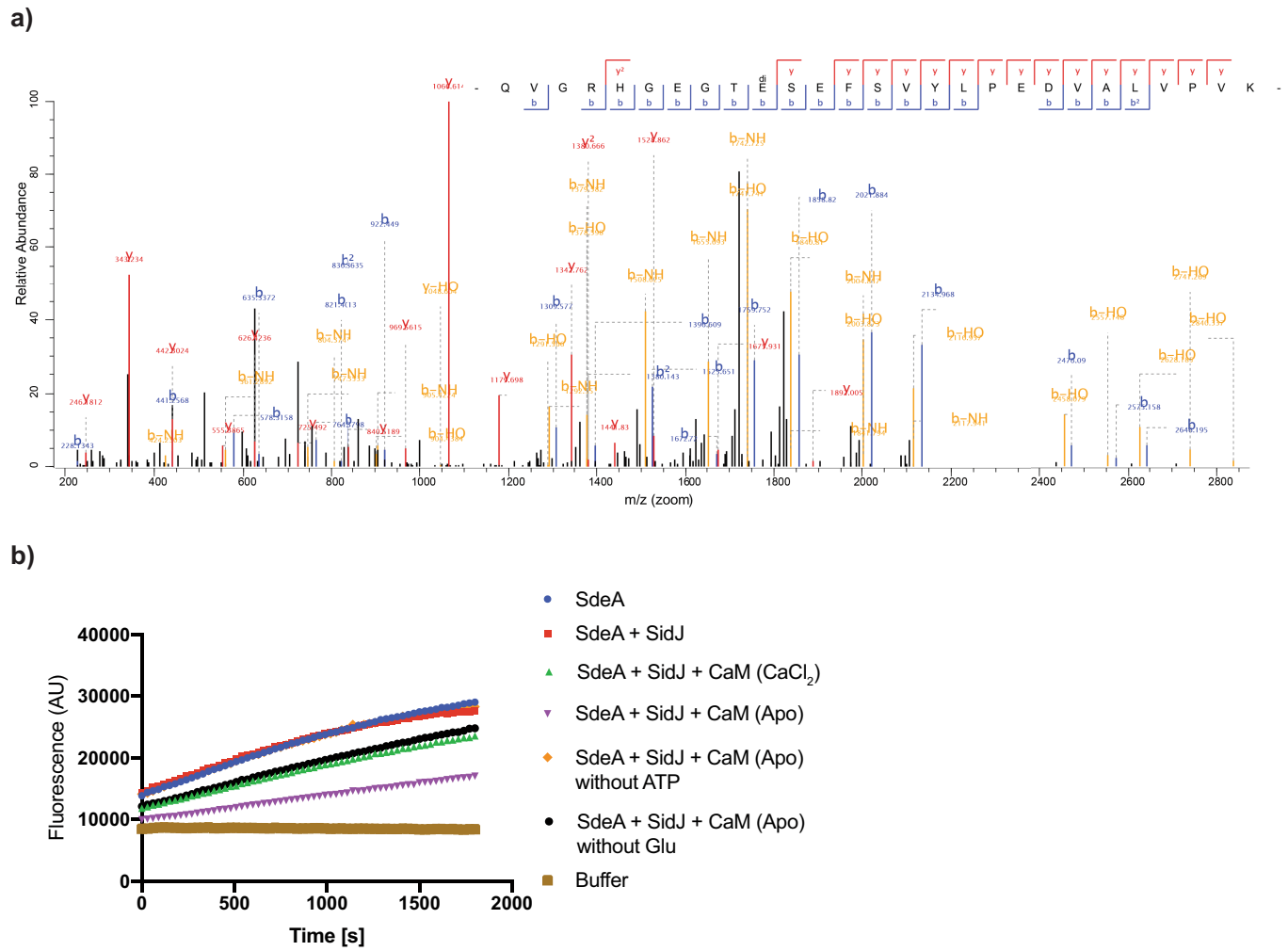


Extended Data Fig. 1 | SidJ does not possess intrinsic deubiquitinase activity. a, Genetic locus of *sdeC-orf2-sidj-sdeB-sdeA* in the *Legionella* genome. **b**, Left, GFP–SidJ that was ectopically expressed and purified from HEK293T cells was incubated with canonical HA–ubiquitin chains purified from mammalian cells. The canonical deubiquitinase USP2 was used as a positive control. Right, GFP or GFP–SidJ was incubated with purified SdeA–ubiquitinated Rab33b. The experiment was repeated twice independently, with similar results. **c**, Full-length SidJ was incubated with various substrates modified with canonical ubiquitination or phosphoribosyl-linked ubiquitination, to probe the cleavage activity. USP2 was used as a positive control for cleaving canonical ubiquitin chains. The experiment was repeated twice independently, with similar results. **d**, SidJ purified from *Legionella* was analysed by mass spectrometry, and protein quantification was performed using the MaxQuant iBAQ algorithm. The experiment was repeated twice independently, with similar results.

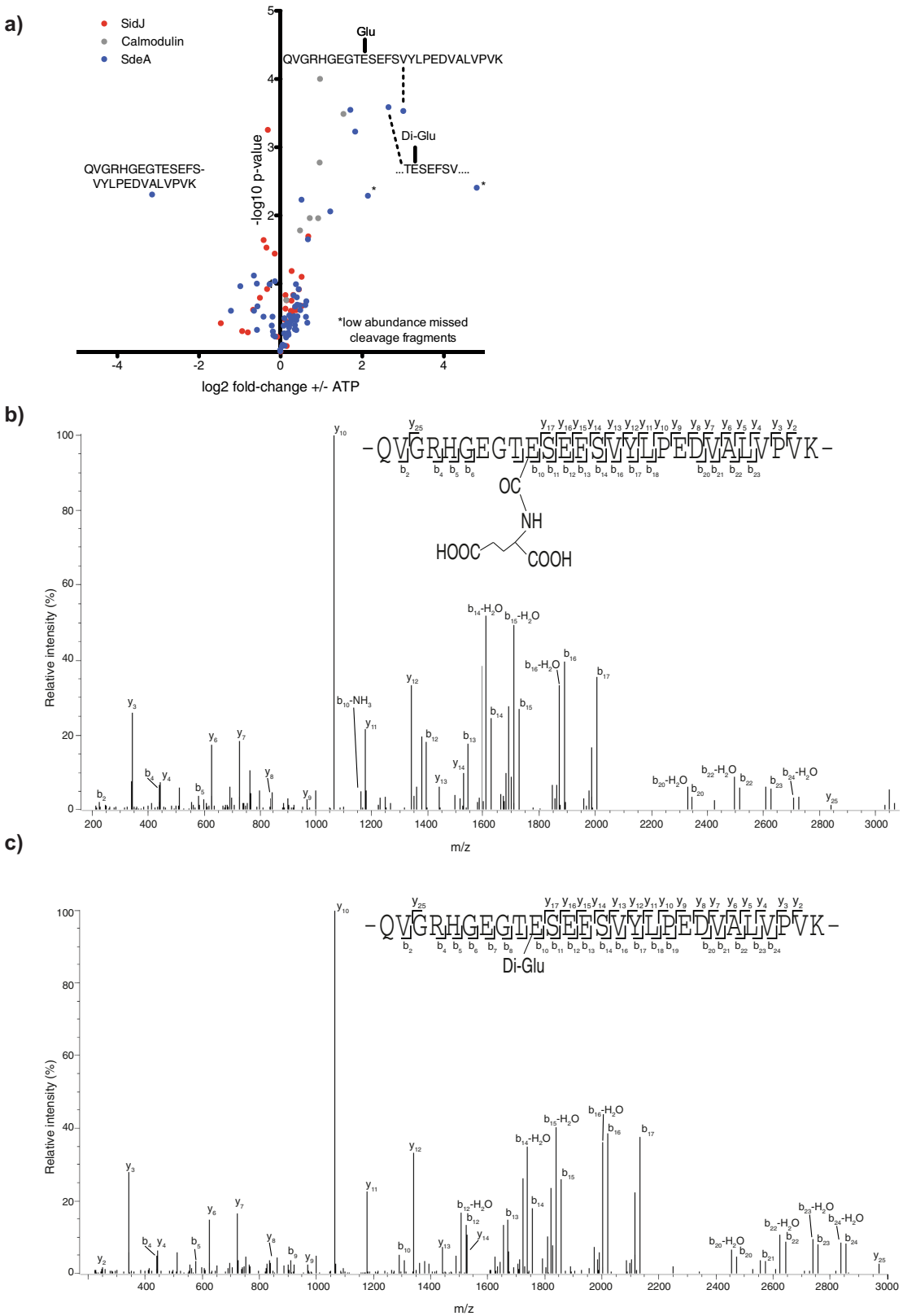


Extended Data Fig. 2 | See next page for caption.

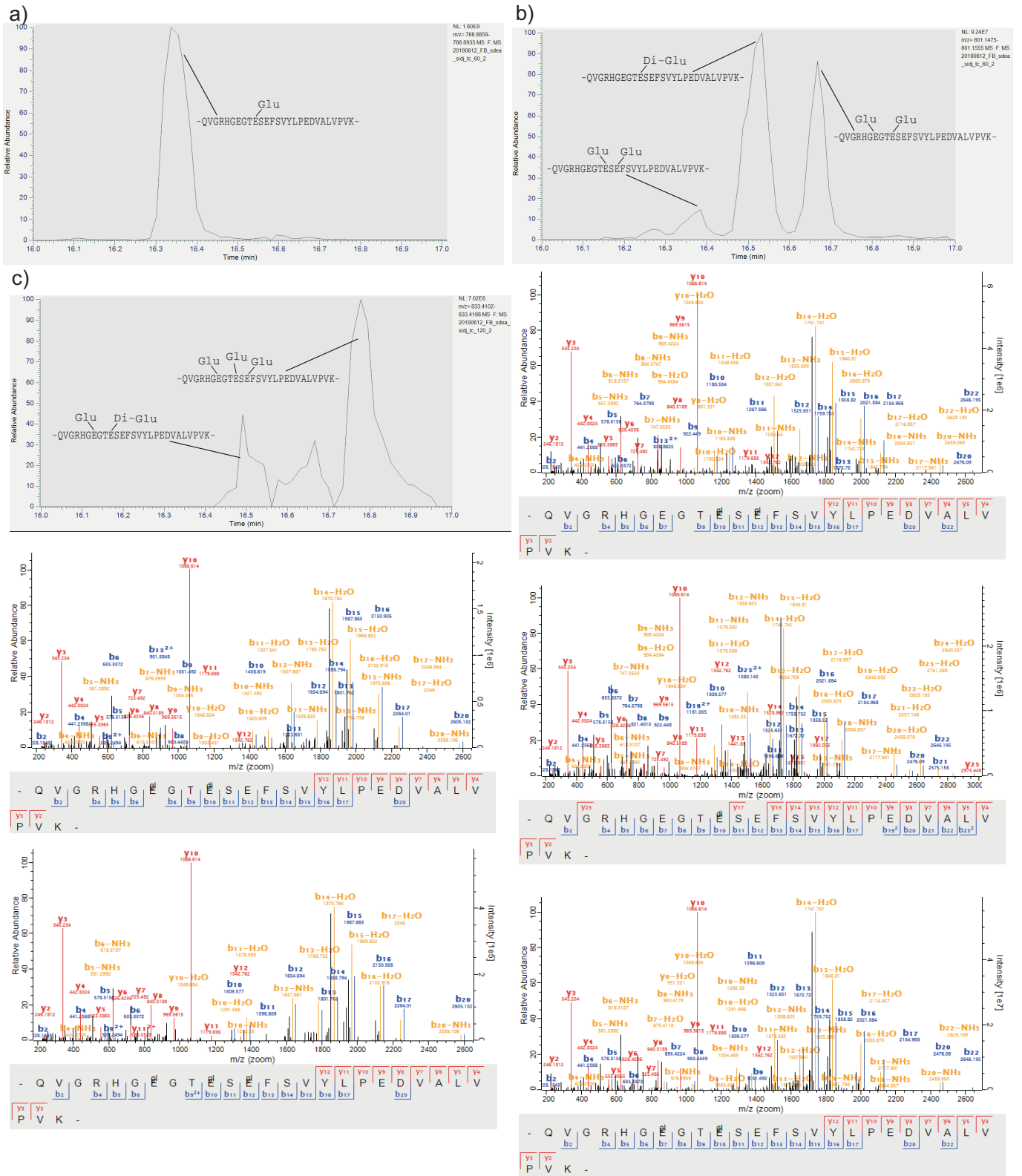
Extended Data Fig. 2 | SidJ binds to CaM. a, GFP and GFP–SidJ were ectopically expressed in HEK293T cells, and immunoprecipitated. The samples were analysed by SDS–PAGE, followed by Coomassie staining. The experiment was repeated twice independently, with similar results. **b**, Purified SdeA from HEK293T cells that express SdeA alone, or in combination with SidJ or SidJ(ΔIQ) (here labelled SidJ(1–819)), was used in ε-NAD^+ hydrolysis assays. The experiment was repeated twice independently, with similar results. **c**, A549 cells infected with wild-type *L. pneumophila* for different time periods were loaded with Fura2AM for 30 min at 37 °C, followed by ratiometric measurement of intracellular Ca^{2+} using a plate reader. Infection with bacteria did not change total Ca^{2+} levels in the cell. $n = 3$ biologically independent experiments, Data points indicate mean and error bars represent s.d. **d**, A549 cells expressing endoplasmic reticulum–GFP and endoplasmic reticulum–cepia were infected with *L. pneumophila* (Ds-Red lp02) followed by time-lapse imaging. Fluorescence intensity of endoplasmic reticulum–cepia at each region is proportional to local Ca^{2+} levels. Endoplasmic reticulum–GFP fluorescence is independent of Ca^{2+} concentration. The endoplasmic reticulum has a heterogenous and dynamic distribution of Ca^{2+}. The bacteria make transient contacts with the endoplasmic reticulum, and may be influenced by local Ca^{2+} fluxes in the cell. Endoplasmic reticulum–cepia is marked in red, endoplasmic reticulum–GFP in green and bacteria are marked by white dotted lines. Time-lapse images were taken at 1-s intervals for 2 min. Images shown in the montage are at 10-s intervals. The experiment was repeated three times independently, with similar results. **e**, Crystal structure of SdeA (PDB 5YIM) is shown in cartoon representation, highlighting the missing peptide of SdeA in SidJ-treated samples (shown in red). The solvent-exposed part of this peptide that contains the catalytic glutamate (E860, shown in green) is marked.



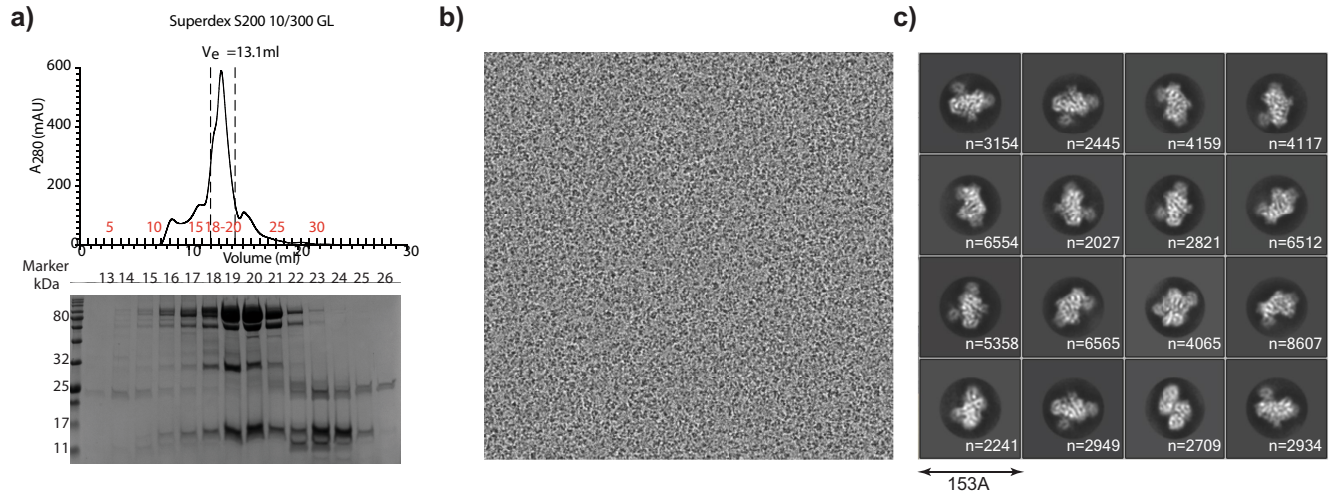
Extended Data Fig. 3 | Mass spectra showing glutamylation of E860 of SdeA. a, Annotated mass spectra for di-glutamylation of E860 of SdeA from samples of immunoprecipitated GFP–SdeA co-expressed with SidJ. **b**, In vitro glutamylation of SdeA was carried in various conditions (as indicated), followed by ε-NAD^+ hydrolysis assays to measure the ubiquitin–ADP ribosylation activity of SdeA. The experiment was repeated three times independently, with similar results.

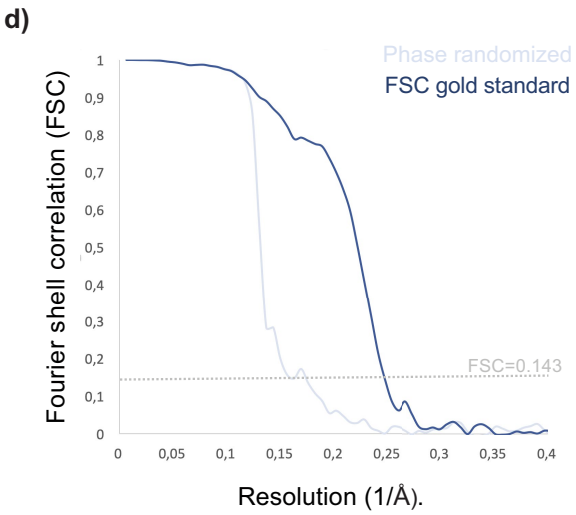


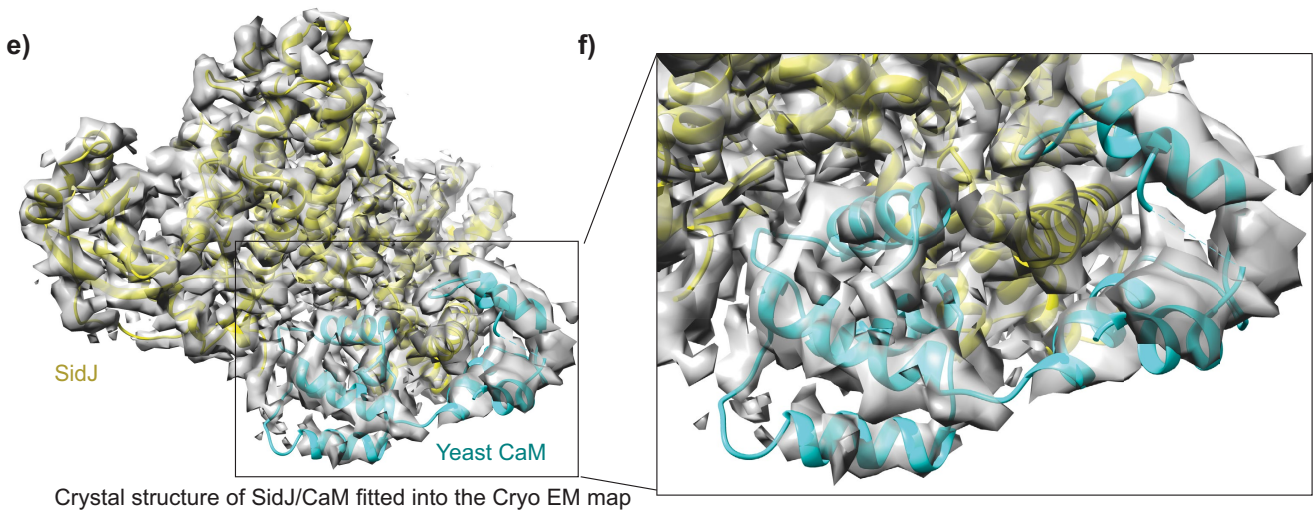
Extended Data Fig. 4 | In vitro glutamylation of SdeA. a, Samples of in vitro glutamylation reactions that contained SdeA and SidJ, with or without ATP, were TMT-labelled and analysed by quantitative mass spectrometry. Mono- and di-glutamylation of the catalytic E860 of SdeA was enriched in samples that contained ATP. In vitro glutamylation was performed in $n = 3$ biologically independent experiments. Samples were labelled with TMT six-plex reagent and analysed in one liquid chromatography–mass spectrometry run. Significant differences between samples were detected by a two-sided Student's *t*-test. **b**, Annotated mass spectra for mono-glutamylation of E860 of SdeA, from SidJ glutamylase in vitro reactions. **c**, Annotated mass spectra for di-glutamylation of E860 of SdeA, from SidJ glutamylase in vitro reactions.



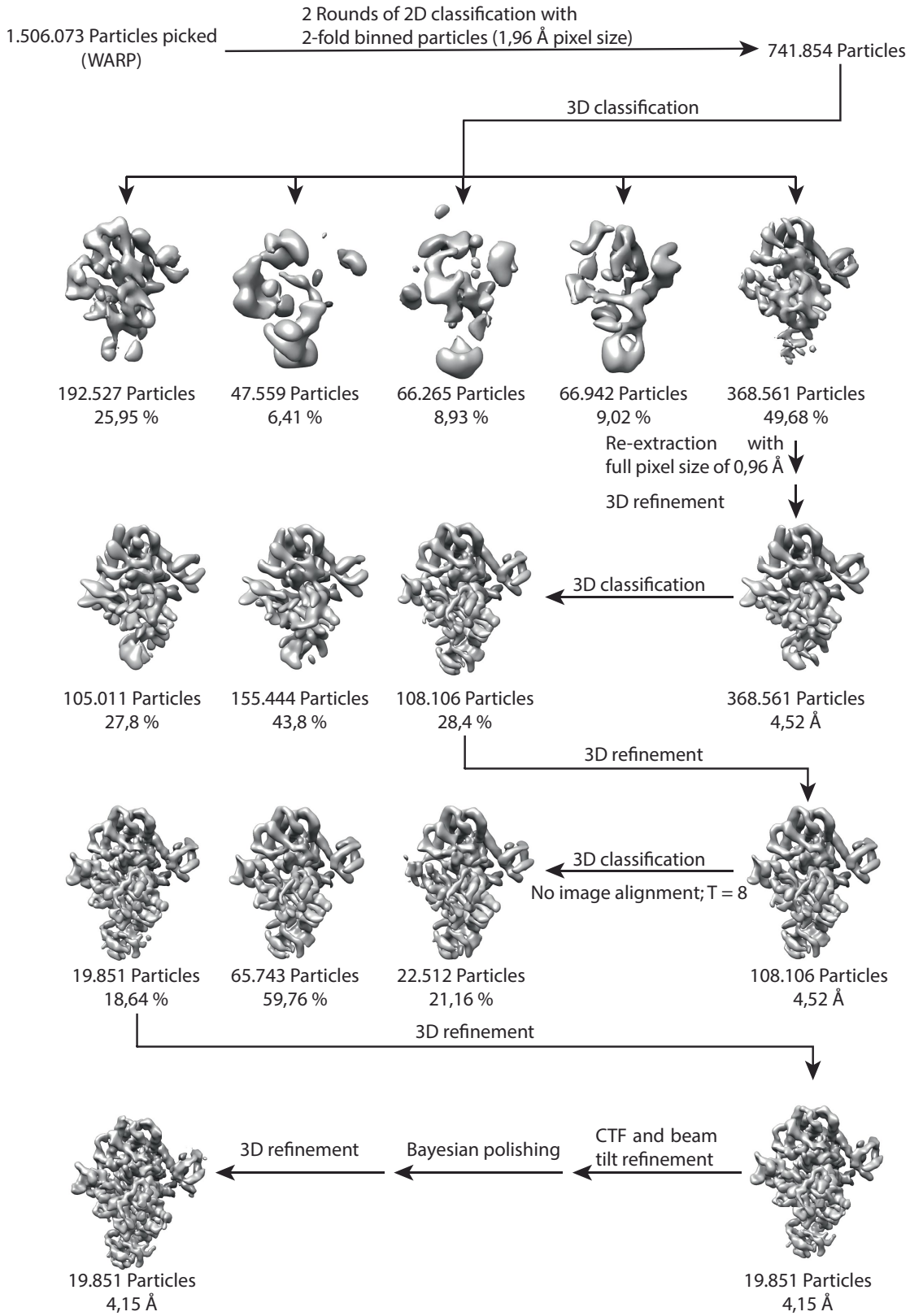
Extended Data Fig. 5 | Analysis of isobaric glutamylated peptide species. a, Extracted ion chromatogram of +4-charged QVGRHGEGTESEFSVYLPEDVALVPVK peptide with +1 glutamate, showing that there are no other co-existing mono-glutamylated versions of the catalytic peptide (besides glutamylation of E860). **b**, Extracted ion chromatogram of the catalytic peptide plus two glutamates (charge +4) is separated into three peaks that could be assigned to di-glutamylation of E860, as well as two mono-glutamylations on the peptide on E860 and E857, and on E860 and E862. Annotated spectra are shown below. **c**, Extracted ion chromatogram of the catalytic peptide plus three glutamates (charge +4) is separated into three different peaks that could be assigned to di-glutamylation of E860, plus mono-glutamylation of E857 and a parallel mono-glutamylation of E857, E860 and E862; a third peak could not be clearly assigned. Annotated spectra of the annotated species are shown below. In **a**–**c**, in vitro glutamylation and label-free liquid chromatography–mass spectrometry analysis were performed in three biologically independent experiments with similar results. Corresponding quantitative information is shown in Fig. 3d.



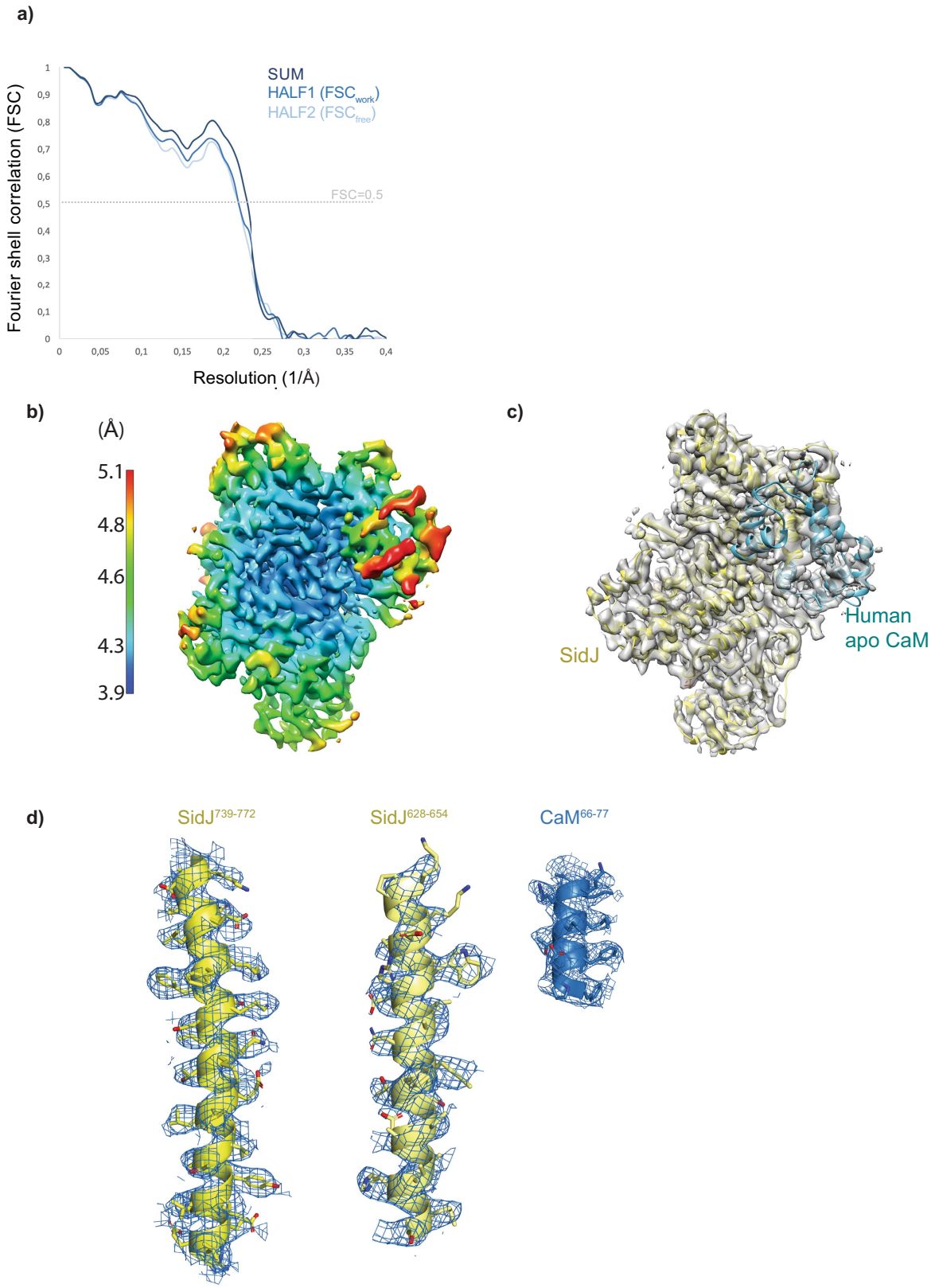




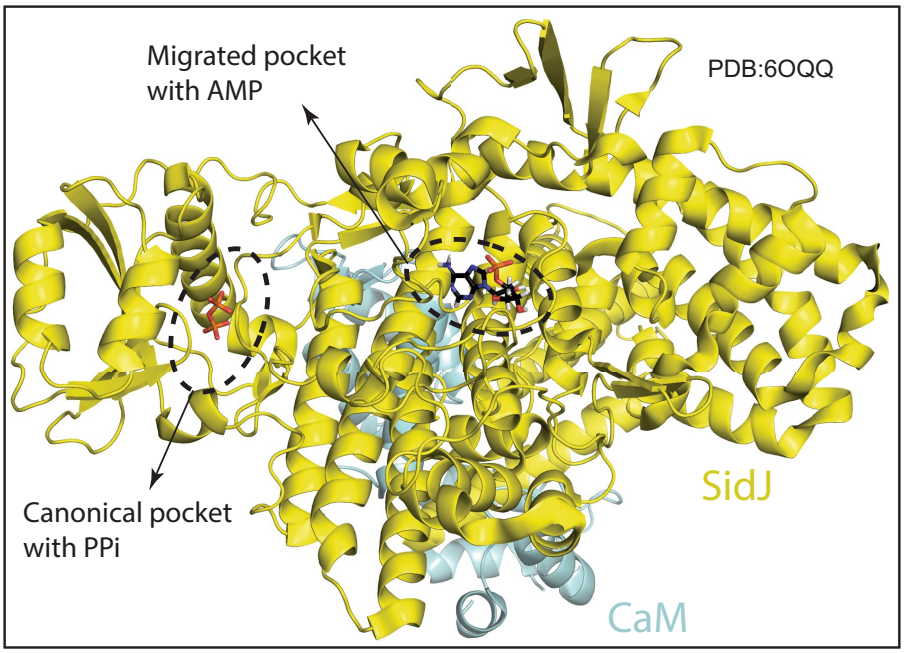
Extended Data Fig. 6 | Cryo-EM data processing and 3D reconstruction. **a**, Size-exclusion profile of SidJ–CaM complex. Elution fractions were analysed by SDS–PAGE. Marked fractions were used for cryo-EM sample preparation. This experiment was repeated three times independently, with similar results. **b**, A representative electron micrograph for the cryo-EM dataset collected. **c**, Reference-free representative 2D class averages of the SidJ–CaM complex. Secondary structure features are visible in projection images. The number (*n*) of particles used to obtain a 2D class average is mentioned in each subpanel. **d**, Gold-standard FSC[30] plot between two independently refined half-maps, $FSC_{0.143} = 4.15$-Å resolution. As expected, FSC between phase-randomized half-maps show a rapid drop of correlation beyond randomization point. **e**, Crystal structure of SidJ–CaM (PDB 6OQQ) is fitted into the cryo-EM 3D reconstruction. **f**, Part of **d**, magnified to highlight the difference between the crystal structure and the cryo-EM map.

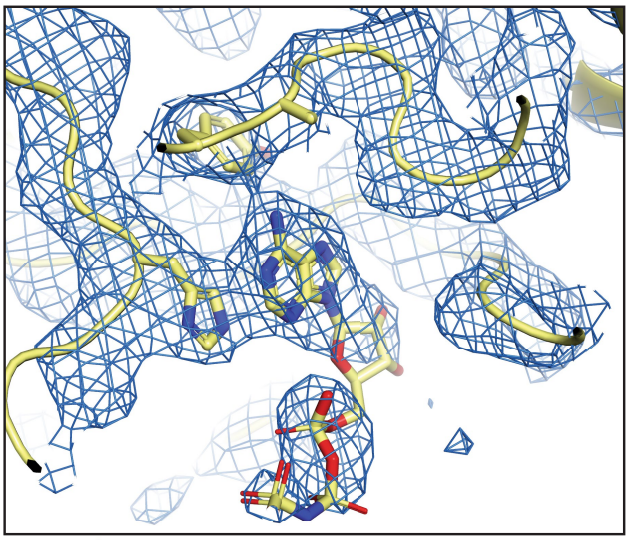


Extended Data Fig. 7 | Cryo-EM single-particle analysis pipeline. Data-processing strategy for cryo-EM of the SidJ–CaM complex. Particle-picking on 2,423 micrographs (using WARP) resulted in the identification of 1,500,000 particles. The particle coordinates were imported into Relion and particles were extracted with a twofold binning factor. After 2D classification and an initial 3D classification, a 3D class with clear secondary structure features and 370,000 particles was identified. The particles of this class were re-extracted with full pixel size and 3D-refined, resulting in a 4.52-Å model. Two additional rounds of 3D classification and 3D refinement improved the resolution of the model to 4.15 Å. Final particle polishing and CTF refinement of the remaining particles did not result in a nominal improvement of the resolution.

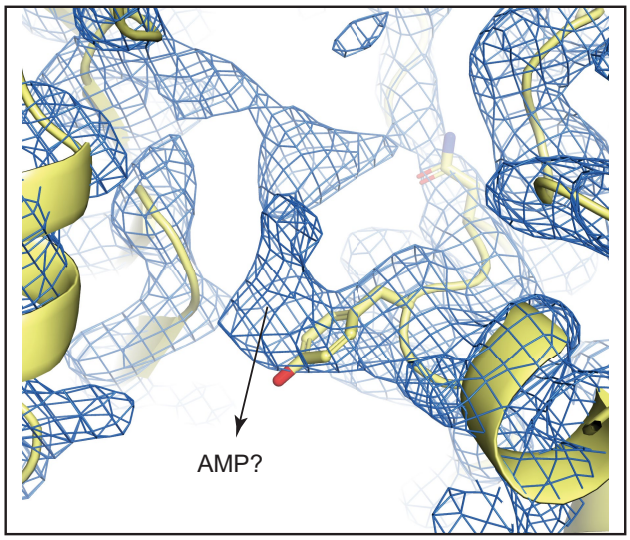


Extended Data Fig. 8 | Cryo-EM model refinement. a, FSC between model and the map and cross-validation of the model fitting. $FSC_{0.5}$ = 4.3 Å for the model versus map (sum). Half-map cross validation procedure[31] does not show overfitting in the refined model. **b**, Local resolution analysis (using Relion) shows variation in the map resolution ranging from 3.9 to 5.1 Å. **c**, Overview of the model fitting into the map in the same orientation as in **b**. **d**, An example of the cryo-EM map quality with the atomic model fitted in, showing clear density for side chains.



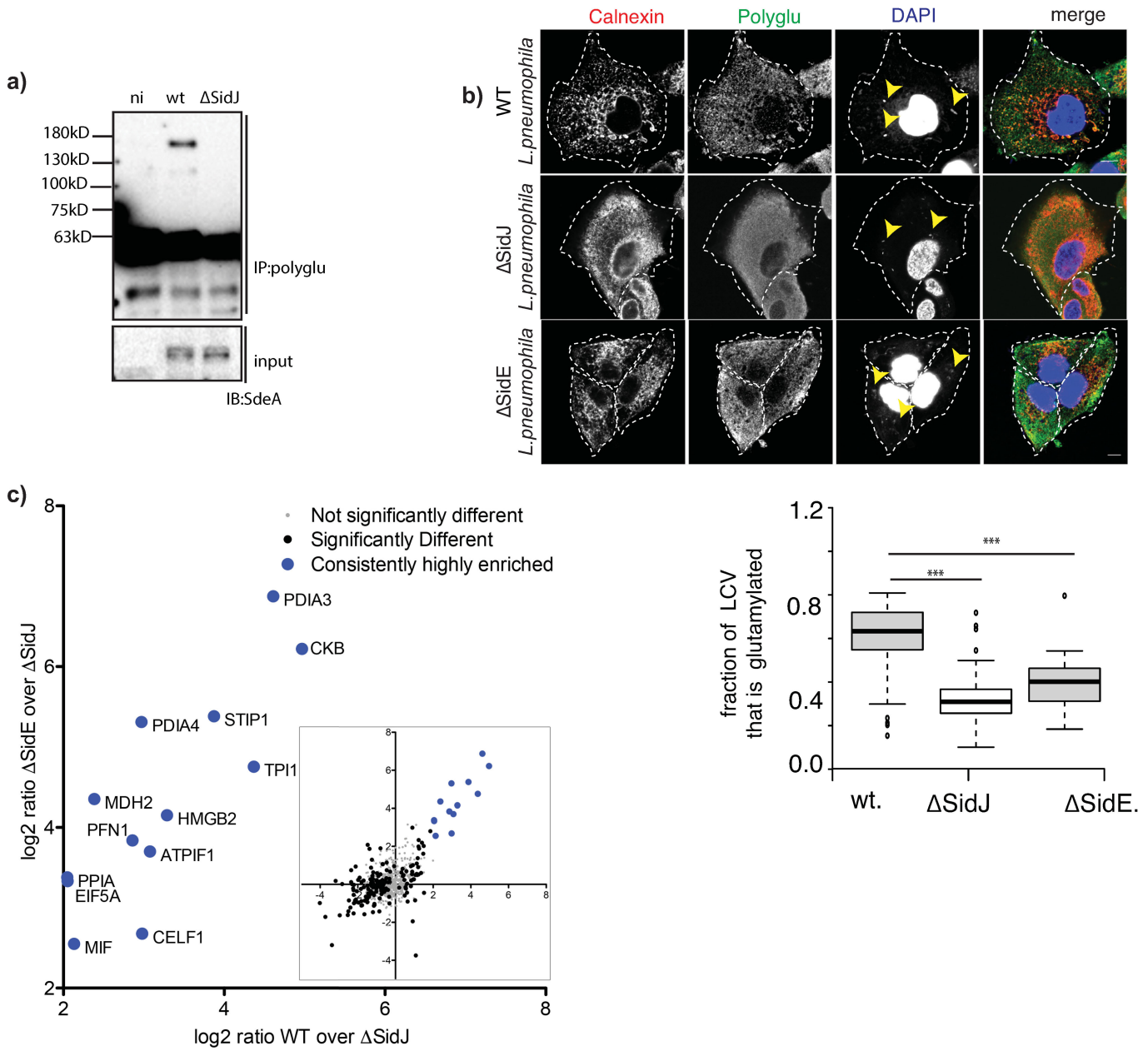


EM map showing electron density in the migrated pocket fitted with AMPPNP



EM map showing unassigned electron density in the canonical pocket

Extended Data Fig. 9 | Ligand-binding sites of SidJ–CaM complex. **a**, Crystal structure of complex between SidJ and yeast CaM (PDB 6OQQ) is shown, marking the two proposed catalytic sites and the bound ligands. **b**, Electron-microscopy map showing cryo-EM density in the migrated pocket fitted with AMPPNP. **c**, Electron-microscopy map showing unassigned cryo-EM density in the canonical pocket.



Extended Data Fig. 10 | SidJ-dependent glutamylation during *Legionella* infection. a, Raw264.7 macrophages were infected with wild-type, *ΔSidJ* or *ΔSidE Legionella* for 3 h. Lysates were used for immunoprecipitation with polyglutamylation antibody, followed by immunoblotting with SdeA. ni, samples that were not infected with bacteria. This experiment was repeated twice independently, with similar results. **b**, A549 cells were infected with different strains of *L. pneumophila* for 3 h. Cells were fixed and immunostained with antibodies against calnexin and polyglutamylation (GT335). DAPI staining marks the nucleus and cytosolic bacteria. Yellow arrows indicate bacteria in infected cells. The ROI is defined as calnexin-stained LCV. ROIs of 80 × 100 μm^2 were chosen in the perinuclear region of cells, followed by quantification of Mander's coefficient (*m*) using the Coloc2 plugin in FIJI. *m* represents the fraction of calnexin-positive LCVs that are also positive for polyglutamylation. Centre lines show the medians; box limits indicate the 25th and 75th percentiles (as determined by R software); whiskers extend 1.5× the interquartile range from the 25th and 75th percentiles; and outliers are represented by dots. Number of ROIs (*n*) = 80 from 30 cells was used for quantification. ***$P < 0.001$ by two-tailed type-3 Student's *t*-test. *P* value (wild type versus *ΔsidJ*) = 6.18×10^{-29}; *P* value (wild type versus *ΔsidE*) = 1.09×10^{-5}. This experiment was repeated twice independently, with similar results. **c**, Glutamylated proteins were isolated from wild-type, *ΔsidE* and *ΔsidJ Legionella* infection experiments using GT335 antibody and quantified using mass spectrometry. Correlation between wild-type versus *ΔsidJ* and *ΔsidE* versus *ΔsidJ* quantifications are plotted (inset), showing the most-correlated proteins in these two quantifications. *Legionella* infection and label-free liquid chromatograph–mass spectrometry analysis was performed with *n* = 3 biologically independent experiments. Significant differences between samples were detected by a corrected, two-sided Student's *t*-test with permutation-based false-discovery rate of 0.05. Proteins were labelled as significant if they were above the false-discovery-rate threshold of 0.05 in at least one comparison (*ΔsidE* and wild-type *Legionella* compared to *ΔsidJ*-infected cells). Proteins with a $\log_2$ ratio above two (mean) in wild-type samples were labelled as highly enriched compared to *ΔsidJ*-infected cells in samples from wild-type and *ΔsidE*-infected cells.

nature research

Corresponding author(s): Sagar Bhogaraju, Ivan Dikic

Last updated by author(s): Jul 5, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Image lab software 5.2.1,
Data analysis	Image lab software 5.2.1 ,Prism5, Adobe Illustrator CS5, MaxQuant 1.6.5 , Perseus 1.6.5, Relion 3.0, WARP 1.0.6, Chimera, Pymol, CCP-EM

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Mass spectrometry data has been deposited in the Proteomics Identification (PRIDE) database with the dataset identifier PXD014362. Cryo-EM structure coordinates are available from the Protein Data Bank (PDB) and the Electron Microscopy Data Bank (EMDB) under accession codes 6S5T and EMD-10100.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample size calculation was done. Sample size was chosen based on the consistency and magnitude of measured differences between groups.
Data exclusions	No data were excluded from analysis.
Replication	All attempts replicated similar results. We have repeated each experiment in the manuscript at least twice to ensure consistent results. We have also provided detailed protocols in the methods section that will help replicate these results for other researchers.
Randomization	No randomization was necessary as various infection samples were recorded and analyzed by a computer software for extracting the significant differences.
Blinding	Blinding was not relevant for the experiments done as various infection samples were imaged and analyzed by a computer software for extracting the significant differences.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	anti-Polyglutamylation Modification (Cat# GT335, Lot#A23941207, Provider:Adipogen technology, 1µg of antibody was used for immunoprecipitation (IP) of 100µg of lysate) Ubiquitin (Cat# 3936S, Lot#14, Provider: Cell signaling Technology, Dilution: 1:1000) Ubiquitin Ubi-1 (Cat# ab7254, Lot#GR3174867-10, Provider: Abcam, Dilution: 1:5000) GFP (B-2) (Cat # sc-9996, Lot# K2217, provider: Santa Cruz Biotechnology,) Anti-tubulin (cat #:2128S, Lot#5, Provider: Cell signaling Technology, Dilution: 1:5000) Anti-HA (cat #:sc-7392 Lot#L1018, Provider:Santa Cruz Biotechnology, Dilution: 1:1000) Anti-CaM (Cat #: 4830S, Lot# 3, Provider: Cell Signaling Technology, Dilution: 1:2000) Anti-GST (Cat #: 2625S, Lot# 8, Provider: Cell Signaling Technology, Dilution: 1:2000) GFP trap beads (Cat #: gta-100 Lot# 90312001A, Provider: ChromoTek, beads were used for IP according to manufacturer's recommendation)
Validation	anti-Polyglutamylation Modification, mAb (GT335), validation statement from the manufacturer: Recognizes the posttranslational modification (poly)glutamylation on proteins. Reacts with polyglutamylated α- and β-tubulin. validation found at provider's website: https://adipogen.com/ag-20b-0020-anti-polyglutamylation-modification-mab-gt335.html/ Ubiquitin from Cell Signaling Technology (Cat:3936S), validation statement from the manufacture: Ubiquitin (P4D1) Mouse mAb detects ubiquitin, polyubiquitin and ubiquitinated proteins. This antibody may cross-react with recombinant NEDD8. Species Reactivity: All Species Expected, Validation found at provider's website: https://www.cellsignal.com/products/primary-antibodies/ubiquitin-p4d1-mouse-mab/3936

Ubiquitin Ubi-1 from abcam (Cat:ab7254), validation statement from the manufacturer:
Tested applications Suitable for: IHC-P, WB, ELISA, ICC/IF, IHC-Fr
Unsuitable for: IP
Species reactivity Reacts with: Mouse, Rat, Chicken, Cow, Human, Arabidopsis thaliana, Caenorhabditis elegans, Drosophila melanogaster, Drosophila C virus
validation found at provider's website: https://www.abcam.com/ubiquitin-antibody-ubi-1-ab7254.html

CS (cell signaling technology)-Ub and abcam-Ub antibodies were validated in Bhogaraju et al., 2106 Cell for the purpose of monitoring SdeA-mediated ubiquitin modification in HEK293T cells.

GFP (B-2) Santa Cruz Biotechnology (sc-9996)
Validation statement from the manufacturer: recommended for detection of GFP and GFP mutant fusion proteins by WB, IP, IF, FCM and ELISA. Validation found at provider's website https://www.scbt.com/scbt/product/gfp-antibody-b-2?productCanUrl=gfp-antibody-b-2&_requestid=3395142

Anti-tubulin
Validation statement from the manufacturer: β-Tubulin (9F3) Rabbit mAb detects endogenous levels of total β-tubulin protein, and does not cross-react with recombinant α-tubulin.
Species Reactivity:
Human, Mouse, Rat, Monkey, Zebrafish, Bovine Validation found at provider's website: https://www.cellsignal.com/products/primary-antibodies/b-tubulin-9f3-rabbit-mab/2128

Anti-HA
Validation statement from the manufacturer: specific to epitope mapping within an internal region of the the influenza hemagglutinin (HA) protein protein, recommended for detection of proteins containing the HA tag by WB, IP, IF, FCM and ELISA
Validation found at provider's website: https://www.scbt.com/scbt/product/ha-probe-antibody-f-7?productCanUrl=ha-probe-antibody-f-7&_requestid=3428134

Anti-CaM
Validation statement from the manufacturer: Calmodulin Antibody detects endogenous levels of total calmodulin protein.
Species Reactivity:
Mouse, Rat
Species predicted to react based on 100% sequence homology:
Human, Monkey, Xenopus, Pig
Validation found at provider's website: https://www.cellsignal.com/products/primary-antibodies/calmodulin-antibody/4830

Anti-GST
Validation statemet from the manufacturer: GST Antibody detects transfected GST fusion proteins.
Validation found at provider's website: https://www.cellsignal.com/products/primary-antibodies/gst-antibody/2622

GFP-Trap beads:
Validation statemet from the manufacturer : GFP-Trap® Agarose is an affinity resin for immunoprecipitation of GFP-fusion proteins.
It consists of a GFP Nanobody/ VHH coupled to agarose beads.
Specificity
GFP, EGFP, CFP, YFP, BFP and many more derivatives

Validation found at provider's website: https://www.chromotek.com/products/detail/product-detail/gfp-trap-agarose/

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	RAW 264.7 cells (ATCC® TIB-71™), A549 cells (ATCC® CCL-185™) and HEK293T (ATCC® CRL-3216™).
Authentication	Cell lines were authenticated using STR DNA profiling.
Mycoplasma contamination	All the cell lines used tested negative for mycoplasma.
Commonly misidentified lines (See ICLAC register)	The cell lines used in the study are not in the commonly misidentified lines list.

LETTER

https://doi.org/10.1038/s41586-019-1439-1

Regulation of phosphoribosyl ubiquitination by a calmodulin–dependent glutamylase

Ninghai Gan[1,8], Xiangkai Zhen[2,3,8], Yao Liu[1,8], Xiaolong Xu[2,8], Chunlin He[4], Jiazhang Qiu[5], Yancheng Liu[1], Grant M. Fujimoto[6], Ernesto S. Nakayasu[6], Biao Zhou[2], Lan Zhao[2], Kedar Puvar[7], Chittaranjan Das[7], Songying Ouyang[2,3]* & Zhao-Qing Luo[1]*

The bacterial pathogen *Legionella pneumophila* creates an intracellular niche permissive for its replication by extensively modulating host-cell functions using hundreds of effector proteins delivered by its Dot/Icm secretion system[1]. Among these, members of the SidE family (SidEs) regulate several cellular processes through a unique phosphoribosyl ubiquitination mechanism that bypasses the canonical ubiquitination machinery[2–4]. The activity of SidEs is regulated by another Dot/Icm effector known as SidJ[5]; however, the mechanism of this regulation is not completely understood[6,7]. Here we demonstrate that SidJ inhibits the activity of SidEs by inducing the covalent attachment of glutamate moieties to SdeA—a member of the SidE family—at E860, one of the catalytic residues that is required for the mono-ADP-ribosyltransferase activity involved in ubiquitin activation[2]. This inhibition by SidJ is spatially restricted in host cells because its activity requires the eukaryote-specific protein calmodulin (CaM). We solved a structure of SidJ–CaM in complex with AMP and found that the ATP used in this reaction is cleaved at the α-phosphate position by SidJ, which—in the absence of glutamate or modifiable SdeA—undergoes self-AMPylation. Our results reveal a mechanism of regulation in bacterial pathogenicity in which a glutamylation reaction that inhibits the activity of virulence factors is activated by host-factor-dependent acyl-adenylation.

Ubiquitination regulates many aspects of immunity, and pathogens have evolved various strategies through which to co-opt the ubiquitin network to promote their virulence[8,9]. One such example is the SidE

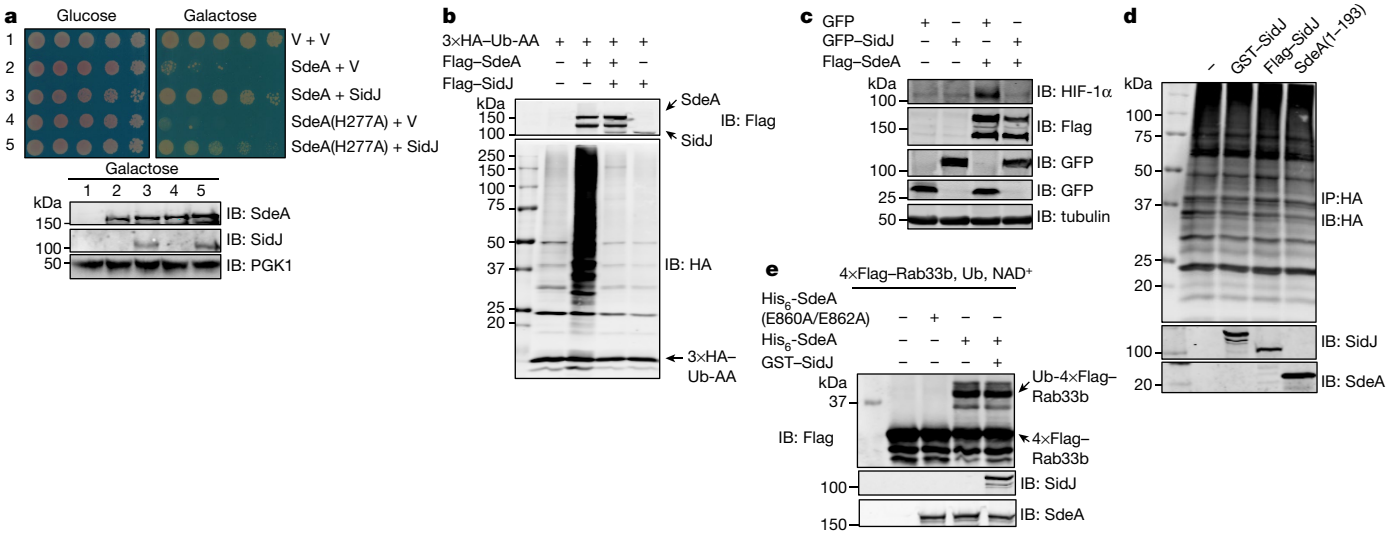


Fig. 1 | SidJ antagonizes the effects of SdeA in eukaryotic cells. a, SidJ suppresses the yeast toxicity of SdeA(H277A). Top, diluted cells from yeast strains inducibly expressing SdeA or SdeA(H277A) that contain the vector (V) or a SidJ construct were spotted onto the indicated media and grown for 2 days. Bottom, the expression of relevant proteins was probed by immunoblotting. The 3-phosphoglyceric phosphokinase-1 (PGK1) was detected as a loading control. **b**, SidJ abrogates SdeA-mediated ubiquitination in mammalian cells. Lysates of HEK293T cells expressing the indicated proteins were detected by immunoblotting with a haemagglutinin (HA)-specific antibody to detect 3×HA–Ub-AA and proteins modified by 3×HA–Ub-AA. The expression of Flag–SdeA and Flag–SidJ was also investigated. **c**, SidJ rescues the degradation of hypoxia-inducible factor 1-α (HIF-1α) that is blocked by SdeA. Lysates of HEK293T cells expressing the indicated proteins were resolved by SDS–PAGE and analysed with antibodies specific for the epitope tags or the relevant proteins. **d**, SidJ from *E. coli* or HEK293T cells cannot deubiquitinate proteins modified by SdeA. Proteins modified by 3×HA–Ub-AA obtained by immunoprecipitation were treated with GST–SidJ from *E. coli*, Flag–SidJ from HEK293T or SdeA(1–193), a truncated form of SdeA containing residues 1–193. Note that none of these proteins caused a reduction in the ubiquitination signals. **e**, GST–SidJ does not inhibit SdeA-induced ubiquitination in vitro. SidJ was co-incubated with SdeA for 2 h at 37 °C and SdeA activity was assayed. A Flag-specific antibody was used to detect modified and unmodified 4×Flag–Rab33b, judging by a shift in its molecular mass. SdeA and SidJ were analysed with specific antibodies. Experiments in each panel were performed independently at least 3 times with similar results.

[1]Purdue Institute for Inflammation, Immunology and Infectious Disease and Department of Biological Sciences, Purdue University, West Lafayette, IN, USA. [2]The Key Laboratory of Innate Immune Biology of Fujian Province, Provincial University Key Laboratory of Cellular Stress Response and Metabolic Regulation, Biomedical Research Center of South China, Key Laboratory of OptoElectronic Science and Technology for Medicine of the Ministry of Education, College of Life Sciences, Fujian Normal University, Fuzhou, China. [3]Laboratory for Marine Biology and Biotechnology, Pilot National Laboratory for Marine Science and Technology (Qingdao), Qingdao, China. [4]Key Laboratory of Organ Regeneration and Transplantation of the Ministry of Education, Department of Respiratory Medicine and Center of Infection and Immunity, The First Hospital of Jilin University, Changchun, China. [5]Key Laboratory of Zoonosis, Ministry of Education, College of Veterinary Medicine, Jilin University, Changchun, China. [6]Biological Science Division, Pacific Northwest National Laboratory, Richland, WA, USA. [7]Department of Chemistry, Purdue University, West Lafayette, IN, USA. [8]These authors contributed equally: Ninghai Gan, Xiangkai Zhen, Yao Liu, Xiaolong Xu. *e-mail: ouyangsy@fjnu.edu.cn; luoz@purdue.edu

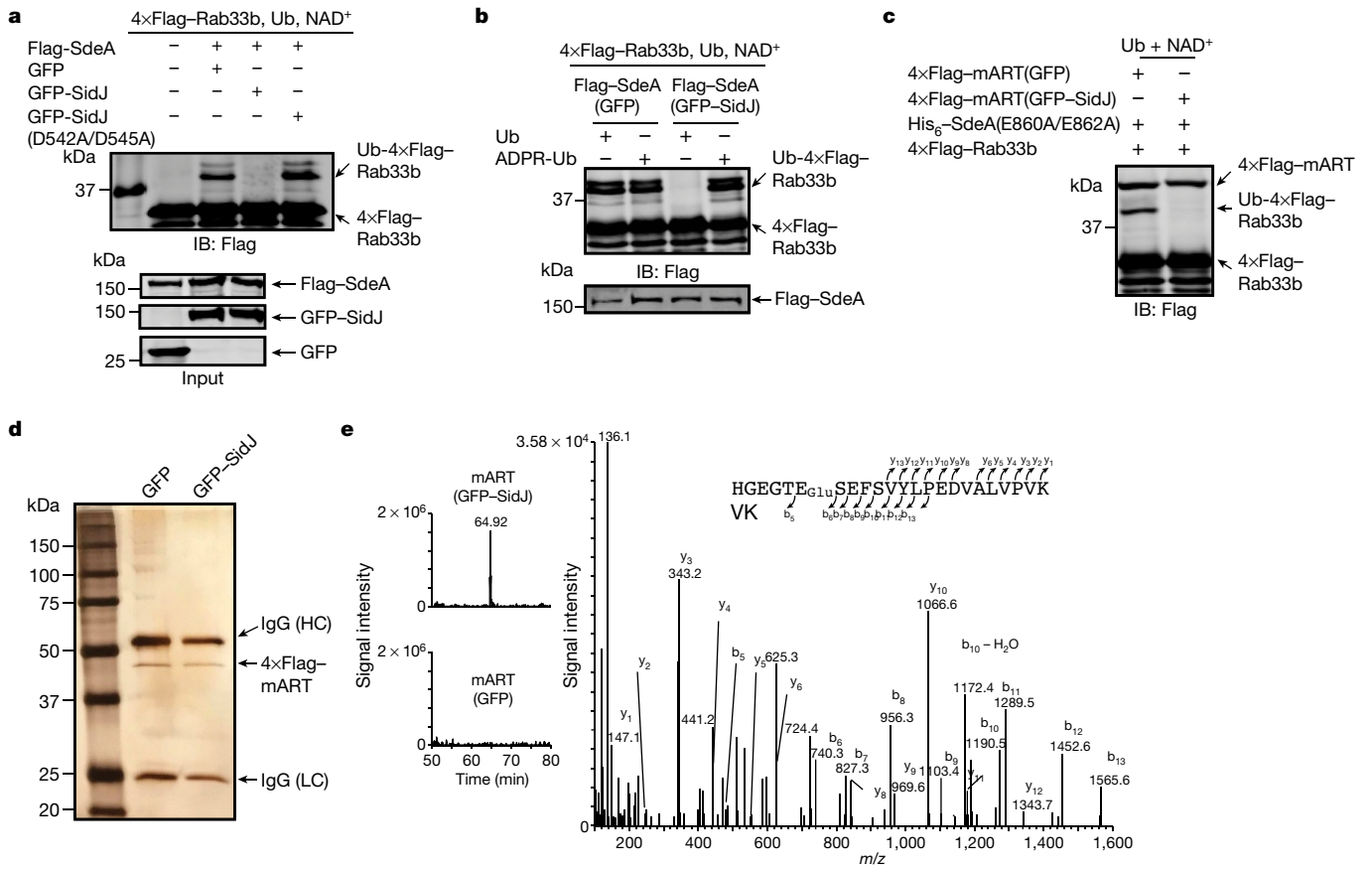


Fig. 2 | SidJ post-translationally modifies SdeA in mammalian cells and inhibits its activity to catalyse the production of ADP-ribosylated ubiquitin. a, Flag–SdeA coexpressed with SidJ fails to modify Rab33b. Flag–SdeA from HEK293T cells coexpressing relevant proteins was used to ubiquitinate 4×Flag–Rab33b. Ub-Rab33b was detected as described in Fig. 1. **b**, Flag–SdeA coexpressed with SidJ retains the ability to ubiquitinate Rab33b with ADPR-Ub. ADPR-Ub or ubiquitin was incubated with Flag–SdeA purified from HEK293T cells coexpressing GFP or GFP–SidJ. NAD^+ was included in reactions that contained ubiquitin. Rab33b modification was detected with a Flag-specific antibody. **c**, SidJ inhibits the mART activity of SdeA. 4×Flag–mART (SdeA(563–910)) purified from HEK293T cells coexpressing GFP or GFP–SidJ was incubated with 4×Flag–Rab33b, ubiquitin, NAD^+ and His_6-SdeA(E/A) for 2 h at 37 °C before the detection of ubiquitination. **d**, **e**, SidJ induces a 129.04-Da post-translational modification on E860 of SdeA. 4×Flag–mART* (**d**) was subject to analysis by mass spectrometry, which identified a posttranslational modification in the fragment –H_{855}GEGTESEFSVYLPEDVALVPVK$_{877}$– (**e**, left). The tandem mass (MS/MS) spectrum shows the fragmentation profile of the modified peptide –H_{855}GEGTE_{Glu}SEFSVYLPEDVALVPVK$_{877}$–, including ions b_5 and b_6 that confirm the modification site at E860 (**e**, right). HC, heavy chain; LC, light chain. The experiment in each panel was repeated three times with similar results.

family effectors from *L. pneumophila*, which ubiquitinate structurally diverse proteins that are associated with the endoplasmic reticulum[2,4]. Ubiquitination by SidEs is initiated by means of ADP-ribosylation at R42 of ubiquitin, which is catalysed by mono-ADP ribosyltransferase (mART)[2]. The activated ADP-ribosylated ubiquitin (ADPR-Ub) is then used by a phosphodiesterase-like domain that is also present in SidEs; this domain ligates phosphoribosylated ubiquitin (PR-Ub) to serine residues of substrate proteins[2,3]. Because both ADPR-Ub and PR-Ub impair the function of eukaryotic cells by inhibiting canonical ubiquitination[3], which is pivotal for bacterial virulence[10], it is likely that there exist factors of either bacterial or host origin that function to prevent potential cellular damage caused by these molecules.

The activity of members of the SidE family—such as SdeA—is regulated by SidJ[5], which is able to suppresses the yeast toxicity of SdeA[6]. SidJ purified from *L. pneumophila* also seems to remove ubiquitin from modified substrates[7]. Despite these observations, questions about the mechanism of action of SidJ remain. For example, an SdeA mutant with a histidine-to-alanine mutation at residue 277 (SdeA(H277A))—which is defective in phosphodiesterase activity—is still toxic to yeast even though it cannot ubiquitinate substrates[3]. However, whether SidJ can suppress its toxicity is unknown. Furthermore, it is not clear why the deubiquitinase activity is observed only in SidJ purified from *L. pneumophila*[7].

We set out to address these questions by constructing a yeast strain that inducibly expressed SdeA(H277A), and found that SidJ effectively suppressed its toxicity (Fig. 1a). SidJ may therefore neutralize the toxicity of ADPR-Ub or target the ADP-ribosylation activity of SdeA. In addition, SidJ substantially reduced protein modification induced by SdeA and effectively relieved SdeA-induced inhibition of the degradation of hypoxia-inducible factor 1α[3] (Fig. 1b, c). However, SidJ that was purified from *Escherichia coli* or from mammalian cells failed to remove ubiquitin from modified proteins, nor did it detectably affect the SdeA-induced ubiquitination of Rab33b (Fig. 1d, e). Together, these results suggest that SidJ affects the function of SdeA, but its activity in cells cannot be recapitulated by biochemical reactions.

Flag-tagged SdeA (Flag–SdeA) that was coexpressed either with GFP or with the SidJ(D542A/D545A) mutant (carrying aspartic-acid-to-alanine mutations at residues 542 and 545), which is defective in suppressing SdeA yeast toxicity[6], was found to robustly modify Rab33b. However, Flag–SdeA obtained from cells coexpressing GFP–SidJ (Flag–SdeA*) failed to ubiquitinate Rab33b (Fig. 2a). We next examined whether SidJ affects the mART activity by carrying out reactions that measure the ability of Flag–SdeA* to use ubiquitin or ADPR-Ub for ubiquitination. Flag–SdeA* lost the ability to catalyse ubiquitination from ubiquitin, but retained the ability to use ADPR-Ub for ubiquitination (Fig. 2b). Consistently, Flag–mART (SdeA residues

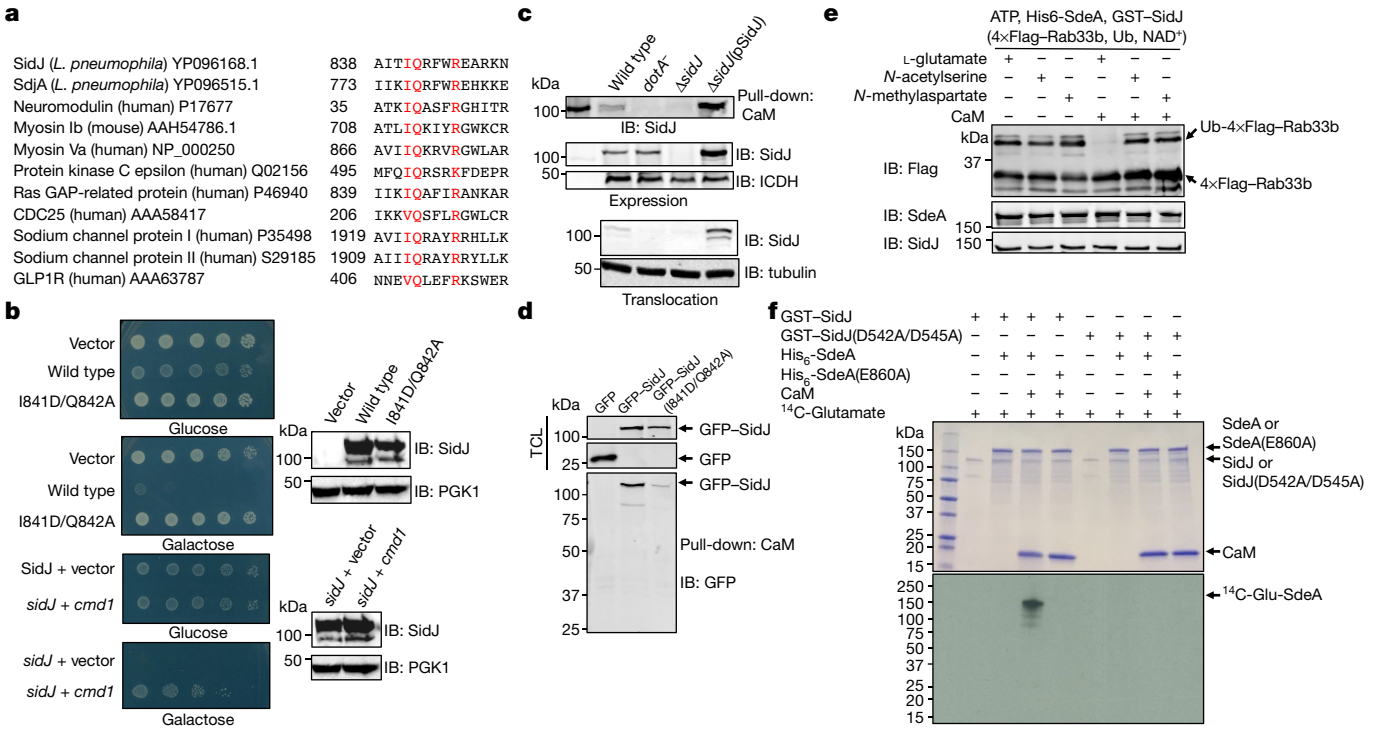


Fig. 3 | Calmodulin is the host cofactor required for the glutamylase activity of SidJ. a, SidJ contains an IQ motif. The list shows the alignment of the IQ domain of SidJ with that of several CaM-binding proteins. Conserved residues are highlighted in red. The accession numbers for each of the proteins (from NCBI databases) are included. **b**, The *cmd1* gene suppresses the yeast toxicity of SidJ. The top two panels show images of serially diluted yeast cells inducibly expressing *sidJ* or its IQ mutant spotted onto the indicated media for 2 days. The lower two panels show the suppression of SidJ toxicity by *cmd1*. The expression of SidJ in each strain was examined and PGK1 was detected as a loading control (right). **c**, **d**, The interactions between SidJ and CaM. Beads coated with CaM were incubated with lysates of macrophages infected with the indicated bacterial strains to analyse its binding to SidJ (**c**, top). SidJ in bacteria (**c**, middle) or translocated into the host cytosol (**c**, bottom) was also examined. The bacterial isocitrate dehydrogenase (ICDH) and tubulin were analysed as loading controls, respectively. Lysates of HEK293T cells transfected to express GFP–SidJ or GFP–SidJ(I841D/Q842A) were incubated with CaM-coated beads (**d**). SidJ or SidJ(I841D/Q842A) bound to CaM was analysed by immunoblotting (bottom). TCL, total cell lysates. **e**, Inhibition of SdeA activity by SidJ requires glutamate and CaM. CaM was added to a subset of a series of reactions containing SdeA, GST–SidJ and L-glutamate, *N*-acetylserine or *N*-methylaspartate. The activity of SdeA was measured by Rab33b ubiquitination. **f**, SidJ is a CaM-dependent glutamylase that modifies SdeA at E860. A series of reactions containing the indicated proteins, ^{14}C-glutamate and ATP were allowed to proceed for 2 h at 37 °C. The incorporation of ^{14}C-glutamate was detected by autoradiography. Data shown in **b**–**f** are one representative of at least three experiments with similar results.

563 to 910, hereafter denoted SdeA(563–910))[11] that was purified from HEK293T cells expressing GFP–SidJ (Flag–mART*) also failed to ubiquitinate Rab33b with the PDE-competent SdeA(E860A/E862A) mutant[3,4] (Fig. 2c). We therefore conclude that SidJ targets the mART activity of SdeA.

Analysis by liquid chromatography coupled with tandem mass spectrometry (LC–MS/MS) identified a mass shift of 129.04 Da (m/z = 129.04, z = 1) on the peptide –H_{855}GEGTESEFSVYLP EDVALVPVK$_{877}$– in Flag–mART* (Fig. 2d, e). The modification—probably the addition of a glutamate—was mapped to E860, one of the catalytic residues of the mART[2] (Fig. 2e). Approximately 93.7% of E860 was modified in samples coexpressed with GFP–SidJ, and a modification of 258.09 Da (m/z = 258.09, z = 1)—presumably diglutamate—was also detected on a small portion of the same peptide (Extended Data Fig. 1). Thus, SidJ may be a glutamylase that ligates one or more glutamate moieties to E860 of SdeA.

We did not detect SidJ activity in reactions containing ATP—the energy source for known glutamylases[12]—and L-glutamate, or its structural isomers *N*-acetylserine and *N*-methylaspartate (Extended Data Fig. 2a). Because the inhibitory effects of SidJ are evident only when it is expressed in mammalian cells, we tested the hypothesis that its activity requires one or more factors of eukaryotic origin by including lysates of *E. coli* or HEK293T cells in the reactions. Lysates of HEK293T cells (native or boiled) caused a decrease in Rab33b modification (Extended Data Fig. 2b), which indicates that one or more heat-stable factors specific to eukaryotic cells are required for the activity of SidJ.

Analysis of the Pfam database[13] of protein families revealed an IQ-like motif, which is involved in CaM binding, located near the carboxyl end of SidJ (Fig. 3a). The yeast toxicity of SidJ[14] was suppressed by mutations in I841 and Q842—two residues in IQ motifs that are important for CaM binding[15]—or by overexpression of the yeast CaM gene *cmd1* (Fig. 3b), thereby validating the nature of the IQ motif. Indeed, binding between SidJ and CaM occurred in cells that were infected with relevant *L. pneumophila* strains or that coexpressed these two proteins, and the IQ motif was found to be required for optimal binding (Fig. 3c, d).

CaM, and SidJ together with L-glutamate—but not the two glutamate isomers—abolished SdeA-mediated ubiquitination (Fig. 3e). Consistent with the heat-insensitivity seen in mammalian cell lysates, boiled CaM was partially active (Extended Data Fig. 2c). Notably, we found that SdeA can be modified by ^{14}C-glutamate only in reactions containing CaM, and that SdeA(E860A) cannot be modified by ^{14}C-glutamate, establishing that E860 is the major modification site (Fig. 3f). Similar to other glutamylases[12], ATP binds SidJ (with a dissociation constant, K_d, of 1.45 μM) and is required for SidJ activity (Extended Data Fig. 2d, e). CaM-dependent inhibition by SidJ was found for all members of the SidE family (Extended Data Fig. 2f). Under our experimental conditions, 0.006 μM and 0.055 μM of CaM was required to activate SidJ and SidJ(I841D/Q842A) respectively; this explained the observation that SidJ(I841D/Q842A) still complemented the phenotype associated with the Δ*sidJ* mutant (Extended Data Fig. 3). Together, these results establish that SidJ is a CaM-dependent

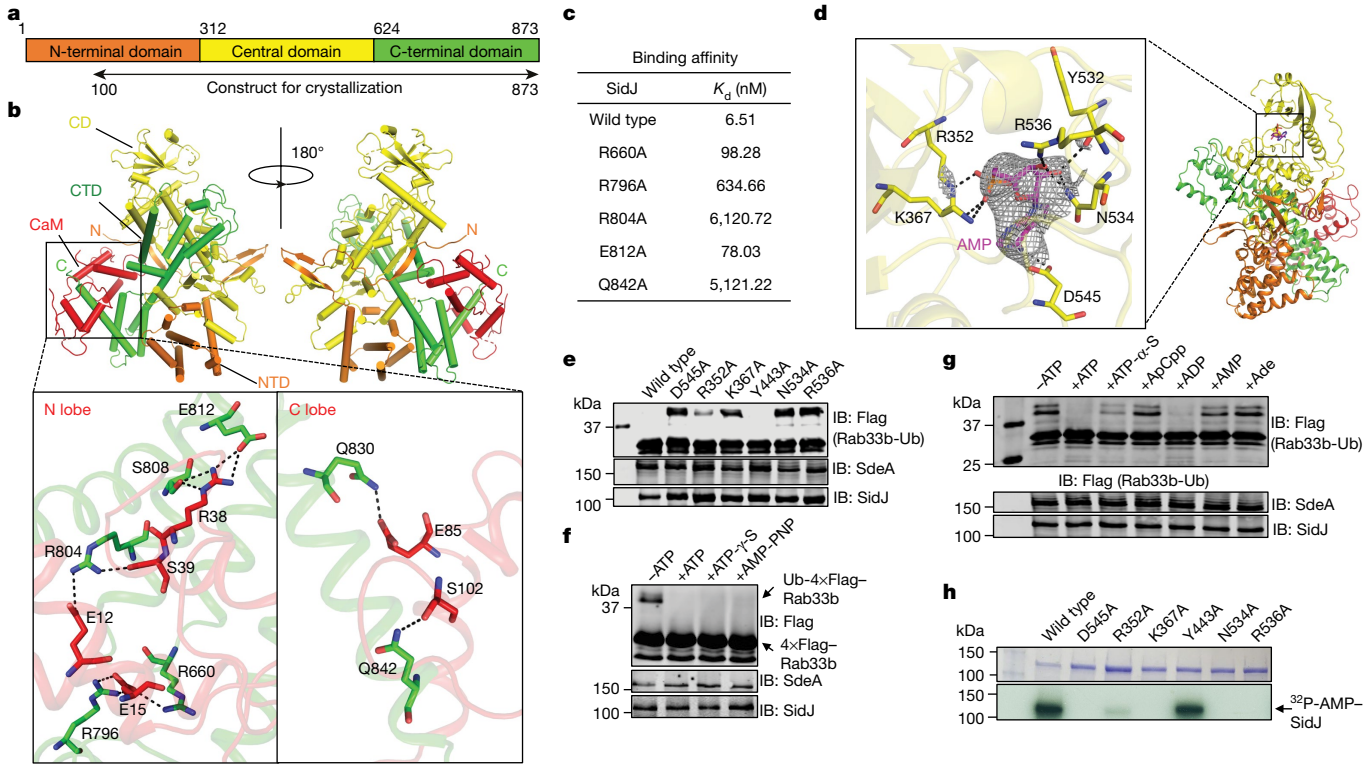


Fig. 4 | Structural analysis of the mechanism of SidJ-catalysed glutamylation. a, The domain organization of SidJ. SidJ consists of the N-terminal domain (orange), the central domain (CD; yellow) and the C-terminal domain (CTD; green). **b**, Ribbon diagram representation of the SidJ–CaM complex. The top panels show the N-terminal domain (NTD; orange), the central domain (CD; yellow) and the C-terminal domain (CTD; green) of SidJ, and CaM (red). The N and C termini of SidJ are labelled with letters. The missing residues are shown as dashed lines. The bottom panels depict interactions between SidJ and the N-lobe and C-lobe of CaM. Residues important for binding are shown as sticks and hydrogen bonds are indicated by dashed lines. **c**, Binding of CaM to SidJ and its mutants. The binding affinity was evaluated using microscale thermophoresis. K_d was calculated by the NanoTemper Analysis 2.2.4 software. Data shown are one representative from three experiments with similar results. **d**, Ribbon representation of the SidJ–CaM–AMP complex. Key SidJ residues involved in AMP binding are shown as yellow sticks, AMP is shown as magenta sticks. Hydrogen bonds are shown as dashed lines. The electron density of a simulated annealing $F_o - F_c$ omit map for AMP is shown, contoured at 3.0σ. **e**, Mutational analysis of residues that are important for the binding of AMP. Each SidJ mutant was incubated with SdeA, ATP, L-glutamate and CaM for 2 h before measuring the ubiquitin ligase activity of SdeA. **f**, **g**, Activation of SidJ by ATP analogues. The indicated compounds were incubated with SdeA, GST–SidJ, L-glutamate and CaM for 2 h at 37 °C before monitoring the activity of SdeA in the ubiquitination of Rab33b. Note that analogues that cannot be hydrolysed at the α site cannot activate SidJ. **h**, The role of residues important for AMP binding in SidJ self-AMPylation. Each SidJ mutant was incubated with ^{32}P-α-ATP, Mg^{2+} and CaM for 2 h at 37 °C and the incorporation of ^{32}P-α-ATP was detected by autoradiography. In **c**, **e**–**h**, data shown are one representative from at least three independent experiments with similar results.

glutamylase that catalyses the ligation of glutamate moieties to E860 of SdeA.

We further investigated the mechanism of the CaM-dependent glutamylase activity of SidJ by structural analysis. A truncated SidJ that lacks the first 99 residues (SidJ(ΔN99)) showed activity that was indistinguishable from that of the full-length protein. Biophysical analysis indicated that it formed a stable heterodimer with CaM at a ratio of 1:1 (Extended Data Fig. 4). We solved a structure at 2.71 Å resolution of the SidJ(ΔN99)–CaM complex, using a 2.95 Å-resolution structure of the SidJ(Se-Met)–CaM derivative as the search model (Extended Data Table 1). In our structure, two SidJ–CaM heterodimeric complexes are found in one asymmetric unit (Extended Data Fig. 5a). Analysis of the intersubunit contacts in the asymmetric unit suggests that the interface between the two SidJ molecules in the structure results from crystal packing. In the complex, SidJ(ΔN99) folds into three distinct domains that we designated as the N-terminal domain, the central domain and the C-terminal domain (Fig. 4a). CaM docks onto the carboxyl end that contains the IQ motif (Fig. 4b). The interface area between SidJ and CaM is about 1,574 Å^2, which accounts for 17.6% of the surface of CaM.

SidJ(ΔN99) interacts extensively with CaM through hydrogen bonds and salt bridges. Specifically, Q830 and Q842 of SidJ engage in hydrogen-bonding interactions with E85 and S102 of the CaM C-lobe, respectively. Other hydrogen bonds include S808(SidJ) and E812(SidJ):R38(CaM), R804(SidJ):S39(CaM), R660(SidJ) and R796(SidJ):E15(CaM) (Fig. 4b). Mutations in these residues reduced the binding affinity of SidJ for CaM (Fig. 4c).

In order to determine the role of ATP in the activity of SidJ, we crystallized the SidJ(ΔN99)–CaM complex in the presence of ATP and obtained a structure at 3.11 Å resolution (Extended Data Table 1). We observed an AMP moiety bound in a pocket formed in the central domain. This domain—along with approximately one hundred additional residues—has been designated as the kinase domain in a recent study[16], in which the same pocket is shown to be occupied by pyrophosphate and Mg^{2+} ions. The AMP moiety—which is probably a product of ATP breakdown, induced by SidJ—is coordinated by R352, K367, Y532, N534, R536 and D545 (Fig. 4d). Substitution of R352, K367, N534, R536 or D545 by alanine abolished the activity of SidJ, whereas a mutation in the distal Y443 had no effect (Fig. 4e). The binding of AMP does not cause obvious conformational changes in the SidJ(ΔN99)–CaM complex (Extended Data Fig. 5b). In our structures, we observed CaM in a relatively closed conformation[17] with one Ca^{2+} coordinated in the EF1 site of CaM (Extended Data Fig. 6a). However, the B-factor of Ca^{2+} was relatively high, indicating partial occupancy of the ion; this is consistent with the partial disorder found in the CaM polypeptide in the crystals. CaM remained active even after dialysis

against EGTA or upon inclusion of this chelator in the reactions (Extended Data Fig. 6b, c).

The presence of AMP in the structure suggests that ATP was cleaved at the α site during the reaction. Indeed, the ATP analogues adenylyl-imidodiphosphate and ATP-γ-S—which cannot be effectively hydrolysed at the γ site—still activated SidJ (Fig. 4f). ADP—but not AMP or adenosine—potently induced SidJ activity; in addition, ATP-α-S, which can be slowly hydrolysed at the α site[18], partially supported SidJ activity. By contrast, ApCpp—which cannot be cleaved at the α-site—failed to detectably activate SidJ (Fig. 4g). We therefore conclude that the SidJ-catalysed reaction involves the cleavage of ATP between the α and β phosphates.

Because SidJ-induced cleavage of ATP is analogous to the reaction involved in AMPylation[19], we thus examined whether SidJ catalyses AMPylation using ^{32}P-α-ATP. Robust self-AMPylation of SidJ was detected in reactions containing CaM; such modification also occurred in glutamylation reactions that lacked glutamate or modifiable SdeA (Extended Data Fig. 7a, b). Furthermore, residues that are important for binding AMP are required for self-modification activity (Fig. 4h). We detected AMP in reactions containing SidJ, CaM and ATP, and the release of AMP was accelerated by SdeA but not by SdeA(E860A) (Extended Data Fig. 7c). We propose a model in which SidJ activates E860 of SdeA by acyl-adenylation, which is followed by nucleophilic attack of the amino group of free glutamate on the activated carbonyl of the unstable E860–AMP intermediate, leading to glutamylation of E860 and the release of AMP (Extended Data Fig. 7d).

Overexpression of SdeA in the Δ*sidJ* mutant severely affected intracellular bacterial replication[6,20], as did expression of SdeA(M408A/L411A), which is defective in substrate recognition[11]. Such defects were rescued by simultaneous expression of SidJ (Extended Data Fig. 8). We attempted to separate the ubiquitin ligase activity from being the substrate for SidJ by constructing the mutant protein SdeA(E860D). SidJ can neither modify this mutant nor suppress its yeast toxicity. Similarly, its ubiquitin ligase activity is insensitive to SidJ. Of most relevance, the inhibition of intracellular growth of the Δ*sidJ* strain by SdeA(E860D) cannot be rescued by coexpressing SidJ (Extended Data Fig. 9).

The AMP-binding site in our structure is essential for the activation step, but it remains unclear how free glutamate is recognized. The E860–AMP intermediate produced at this site may transit to a second nucleotide-binding site in the same domain for glutamylation[16]. It is also not clear how SidJ selectively targets E860 of SdeA, but not nearby E857 and E862, or whether it modifies other proteins as well as SidEs by glutamylation or AMPylation. The glutamylation of SidEs by SidJ expands the strategies used by *L. pneumophila* to ensure balanced modulation of host function[1]. SidJ is a unique glutamylase that bears no similarity to mammalian glutamylases[12,21]. The requirement of CaM for its activity ensures that SidEs will not be inactivated prior to modifying host targets[7]. CaM also activates the oedema factor of *Bacillus anthracis* and CyaA of *Bordetella pertussis*[22,23], both catalysing the synthesis of the important signalling molecule cyclic AMP[24]. Further study of the mechanism of CaM-induced activation of SidJ and the relationship between the AMPylation and glutamylation reactions is likely to reveal insights into the regulation and function of glutamylases.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1439-1.

Received: 24 April 2019; Accepted: 9 July 2019;
Published online 22 July 2019.

1. Qiu, J. & Luo, Z. Q. *Legionella* and *Coxiella* effectors: strength in diversity and activity. *Nat. Rev. Microbiol.* **15**, 591–605 (2017).
2. Qiu, J. et al. Ubiquitination independent of E1 and E2 enzymes by bacterial effectors. *Nature* **533**, 120–124 (2016).
3. Bhogaraju, S. et al. Phosphoribosylation of ubiquitin promotes serine ubiquitination and impairs conventional ubiquitination. *Cell* **167**, 1636–1649 (2016).
4. Kotewicz, K. M. et al. A single *Legionella* effector catalyzes a multistep ubiquitination pathway to rearrange tubular endoplasmic reticulum for replication. *Cell Host Microbe* **21**, 169–181 (2017).
5. Liu, Y. & Luo, Z. Q. The *Legionella* pneumophila effector SidJ is required for efficient recruitment of endoplasmic reticulum proteins to the bacterial phagosome. *Infect. Immun.* **75**, 592–603 (2007).
6. Jeong, K. C., Sexton, J. A. & Vogel, J. P. Spatiotemporal regulation of a *Legionella* pneumophila T4SS substrate by the metaeffector SidJ. *PLoS Pathog.* **11**, e1004695 (2015).
7. Qiu, J. et al. A unique deubiquitinase that deconjugates phosphoribosyl-linked protein ubiquitination. *Cell Res.* **27**, 865–881 (2017).
8. Lin, Y. H. & Machner, M. P. Exploitation of the host cell ubiquitin machinery by microbial effector proteins. *J. Cell Sci.* **130**, 1985–1996 (2017).
9. Song, L. & Luo, Z. Q. Post-translational regulation of ubiquitin signaling. *J. Cell Biol.* **218**, 1776–1786 (2019).
10. Dorer, M. S., Kirton, D., Bader, J. S. & Isberg, R. R. RNA interference analysis of *Legionella* in *Drosophila* cells: exploitation of early secretory apparatus dynamics. *PLoS Pathog.* **2**, e34 (2006).
11. Kalayil, S. et al. Insights into catalysis and function of phosphoribosyl-linked serine ubiquitination. *Nature* **557**, 734–738 (2018).
12. Eddé, B. et al. Posttranslational glutamylation of alpha-tubulin. *Science* **247**, 83–85 (1990).
13. Finn, R. D. et al. The Pfam protein families database. *Nucleic Acids Res.* **38**, D211–D222 (2010).
14. Heidtman, M., Chen, E. J., Moy, M. Y. & Isberg, R. R. Large-scale identification of *Legionella pneumophila* Dot/Icm substrates that modulate host cell vesicle trafficking pathways. *Cell. Microbiol.* **11**, 230–248 (2009).
15. Rhoads, A. R. & Friedberg, F. Sequence motifs for calmodulin recognition. *FASEB J.* **11**, 331–340 (1997).
16. Black, M. H. et al. Bacterial pseudokinase catalyzes protein polyglutamylation to inhibit the SidE-family ubiquitin ligases. *Science* **364**, 787–792 (2019).
17. Van Petegem, F., Chatelain, F. C. & Minor, D. L., Jr. Insights into voltage-gated calcium channel regulation from the structure of the $Ca_V1.2$ IQ domain-Ca^{2+}/calmodulin complex. *Nat. Struct. Mol. Biol.* **12**, 1108–1115 (2005).
18. Bagshaw, C. ATP analogues at a glance. *J. Cell Sci.* **114**, 459–460 (2001).
19. Casey, A. K. & Orth, K. Enzymes involved in AMPylation and deAMPylation. *Chem. Rev.* **118**, 1199–1215 (2018).
20. Luo, Z. Q. & Isberg, R. R. Multiple substrates of the *Legionella pneumophila* Dot/Icm system identified by interbacterial protein transfer. *Proc. Natl Acad. Sci. USA* **101**, 841–846 (2004).
21. O'Hagan, R. et al. Glutamylation regulates transport, specializes function, and sculpts the structure of cilia. *Curr. Biol.* **27**, 3430–3441 (2017).
22. Drum, C. L. et al. Structural basis for the activation of anthrax adenylyl cyclase exotoxin by calmodulin. *Nature* **415**, 396–402 (2002).
23. Guo, Q. et al. Structural basis for the interaction of *Bordetella pertussis* adenylyl cyclase toxin with calmodulin. *EMBO J.* **24**, 3190–3201 (2005).
24. Gancedo, J. M. Biological roles of cAMP: variations on a theme in the different kingdoms of life. *Biol. Rev. Camb. Philos. Soc.* **88**, 645–668 (2013).

METHODS

Media, bacteria strains, plasmid constructions and cell lines. *L. pneumophila* strains used in this study were derivatives of the Philadelphia 1 strain Lp02[25] and were grown and maintained on CYE plates or in ACES buffered yeast extract (AYE) broth as previously described[25]. The *sidJ* in-frame deletion strain has been previously described[5]. *sidJ* and *sdeA* genes and their mutants were cloned into pZLQ[26] or pZL507[27] for complementation. The *E. coli* strains XL1-Blue and BL21(DE3) were used for expression and purification of all the recombinant proteins used in this study. *E. coli* strains were grown in LB medium. Genes for protein purifications were cloned into pQE30 (Qiagen), pGEX-6P-1 (Amersham) and pET-21a (Novagen) for expression. For ectopic expression of proteins in mammalian cells, genes were inserted into the 4×Flag CMV vector[2] or the 3×HAcDNA3.1 vector[28]. HEK293T cells were cultured in Dulbecco's modified minimal Eagle's medium (DMEM) supplemented with 10% FBS. U937 cells were cultured in RPMI 1640 medium supplemented with 10% FBS. The yeast strains BY4741 and W303 were used for toxicity assays. Yeast strains were cultured in YPD media containing yeast extract, peptone and glucose, or SD minimal media containing yeast nitrogen base, glucose and amino acid drop-out mix for selection of transformed plasmids. For GAL1 promoter induction, 2% galactose was used to replace 2% glucose as the sole carbon source in minimal media. To examine the yeast toxicity of SidJ and its mutants, each allele was cloned into pYES1NTA (Invitrogen), which contains GAL1 promoter for inducible expression in yeast[14]. The *cmd1* gene was cloned into p415ADH[29] for expression in yeast. For the suppression of the yeast toxicity of SdeA by SidJ, *sdeA* and its mutants were expressed from pYES1NTA and *sidJ* was expressed from p425GPD[29]. All mammalian cell lines were regularly checked for potential mycoplasma contamination by the universal mycoplasma detection kit from ATCC (30-1012K).

Transfection, infection, immunoprecipitation. Lipofectamine 3000 (Thermo Fisher Scientific) was used to transfect HEK293T cells grown to about 70% confluence. Different plasmids were transfected into HEK293T cells. Transfected cells were collected and lysed with the radioimmunoprecipitation assay buffer (RIPA buffer, Thermo Fisher Scientific) 16–18 h after transfection. Cells infected with indicated bacterial strains were similarly processed for immunoprecipitation. When needed, immunoprecipitation was performed with lysates of transfected cells using agarose beads coated with HA-specific antibody (Sigma-Aldrich, A2095), Flag-specific antibody (Sigma-Aldrich, F1804), or CaM (Sigma-Aldrich, A6112) at 4 °C for 4 h. Beads were washed with pre-cold RIPA buffer or respective reaction buffers three times. Samples were resolved by SDS–PAGE and followed by immunoblotting analysis with the specific antibodies, or silver staining following the manufacturer's protocols (Sigma-Aldrich, PROTSIL1).

For infection experiments, *L. pneumophila* strains were grown to the post-exponential phase (optical density at 600 nm, OD_{600}, of 3.3–3.8) in AYE broth. When necessary, complementation strains were induced with 0.2 mM IPTG for 3 h at 37 °C before infection. U937 cells were infected with *L. pneumophila* strains correspondingly. Cells were collected and lysed with 0.2% saponin on ice for 30 min. Cell lysates were resolved by SDS–PAGE and followed by immunoblotting analysis with the specific antibodies. *L. pneumophila* bacteria lysates were resolved by SDS–PAGE followed by immunoblotting with the SidJ- and SdeA-specific antibodies to examine the expression of SidJ and SdeA, and isocitrate dehydrogenase (ICDH) was added as a loading control with the antibodies as previously described[27].

For intracellular growth in *Acanthamoeba castellanii* cells, infection was performed at a multiplicity of infection (MOI) of 0.05 and the total bacterial counts were determined at 24-h intervals as previously described[20]. *A. castellanii* was maintained in HL5 medium. For infection, HL5 medium was replaced by MB medium with 1 mM IPTG added to overexpress the indicated proteins.

Protein purification. Overnight *E. coli* cultures (10 ml) were transferred to 400 ml LB medium supplemented with 100 μg ml^{-1} of ampicillin and the cultures were grown to an OD_{600} of 0.6–0.8 before induction with 0.2 mM IPTG. Cultures were further incubated at 18 °C overnight. Bacteria were collected by centrifugation at 4,000*g* for 10 min, and were lysed by sonication in 30 ml PBS. Bacteria lysates were centrifuged twice at 18,000*g* at 4 °C for 30 min to remove insoluble fractions and unbroken cells. The supernatant containing recombinant proteins was incubated with 1 ml Ni^{2+}-NTA beads (Qiagen) or glutathione agarose beads (Pierce) at 4 °C for 2 h with agitation. Ni^{2+}-NTA beads with bound proteins were washed with PBS buffer containing 20 mM imidazole three times, using 30 times the column volume each time. Proteins were eluted with PBS containing 300 mM imidazole. Glutathione agarose beads were washed with a Tris buffer (50 mM Tris-HCl (pH 8.0)) and eluted with 10 mM reduced glutathione in the same buffer. Proteins were dialysed in buffer containing 25 mM Tris-HCl (pH 7.5), 150 mM NaCl and 1 mM dithiothreitol (DTT) for 16–18 h. The native SidJ(ΔN99) was purified using the same protocol, and the protocol used to purify CaM was similar to this but with the addition of 2 mM $CaCl_2$ and 10% glycol. For crystallization, the SidJ–CaM complex was formed by mixing these two proteins in 20 mM Tris-HCl (pH 8.0), 150 mM NaCl and 2 mM $CaCl_2$.

For protein purification from mammalian cells, HEK293T cells were transfected with corresponding plasmids to express Flag-tagged proteins. Cells were lysed with RIPA buffer, and subject to immunoprecipitation with beads coated in Flag-specific antibody. Proteins were then eluted from the beads by using 3×Flag peptides following the manufacturer's protocol (Sigma-Aldrich, F4799).

Crystallization. The purity of SidJ(ΔN99)–CaM was around 95% as assessed by SDS–PAGE, and initial crystallization screens of native SidJ–CaM were conducted by sitting-drop vapour diffusion using commercial crystallization screens. The protein concentration used for crystallization was 5–7 mg ml^{-1}. Hampton Research kits were used in the sitting-drop vapour diffusion method to obtain preliminary crystallization conditions at 16 °C. Crystallization drops contained 0.5 μl of the protein solution mixed with 0.5 μl of reservoir solution. Diffraction-quality crystals of SidJ(ΔN99)–CaM and its complex with ATP (SidJ(ΔN99)–CaM–ATP) were grown in the presence of 0.1 M HEPES (pH 6.5–7.5), 20% (v/v) PEG 4000, and 0.2 M NaCl. To solve the phase problem, Se-Met was incorporated into SidJ and the SidJ(Se-Met) was purified similarly to native SidJ except with the addition of 5 mM DTT to the buffer during the purification process. The concentration of SidJ(Se-Met)–CaM used for crystallization was also around 7 mg ml^{-1}. Diffraction quality crystals of SidJ(Se-Met)–CaM were grown and optimized under the same conditions. All crystals were flash-frozen in liquid nitrogen, with the addition of 20–25% (v/v) glycerol as a cryoprotectant.

Data collection and structure determination. X-ray diffraction for SidJ(ΔN99/Se-Met)–CaM, native SidJ(ΔN99)–CaM and SidJ(ΔN99)–CaM–ATP were collected at beamline BL-17U1 of the Shanghai Synchrotron Radiation Facility. All data were indexed and scaled using HKL2000 software[30]. The initial phase of SidJ(ΔN99)–CaM was determined by using the single-wavelength anomalous dispersion phasing method. Phases were calculated using AutoSol implemented in PHENIX[31]. AutoBuild in PHENIX was used to automatically build the atom model. Molecular replacement was then performed with this model as a template to determine the structure of other complexes. After several rounds of positional and B-factor refinement using phenix.refine with TLS parameters alternated with manual model revision using Coot[32], the quality of final models was checked using the PROCHECK program (https://www.ebi.ac.uk/thornton-srv/software/PROCHECK). The quality of the final model was validated with MolProbity[33]. Structures were analysed with PDBePISA (Protein Interfaces, Surfaces, and Assemblies)[34], Dali (http://ekhidna2.biocenter.helsinki.fi/dali), and details of the data collection and refinement statistics are given in Extended Data Table 1. All of the figures showing structures were prepared with PyMOL (http://www.pymol.org). In the final models, the model for the SidJ(ΔN99/Se-Met)–CaM complex contained 91.05%, 8.58% and 0.16% in the favoured, allowed and outlier regions of the Ramachandran plot, respectively. The model for the SidJ(ΔN99)–CaM complex contained 93.25%, 6.63% and 0.10% in the favoured, allowed and outlier regions of the Ramachandran plot, respectively. The final model for the SidJ(ΔN99)–CaM–AMP complex contained 92.10%, 7.76% and 0.00% in the favoured, allowed and outliers regions of the Ramachandran plot, respectively.

Analytic ultracentrifugation. Sedimentation velocity experiments were used to assess the molecular size of the SidJ(ΔN99)–CaM complex at 20 °C on a Beckman XL-A analytical ultracentrifuge equipped with absorbance optics and an An60 Ti rotor (Beckman Coulter). Samples were diluted to an optical density at 280 nm (OD_{280}) of 1 in a 1.2-cm path length. The rotor speed was set to 72,500*g* for all samples. The sedimentation coefficient was obtained using the c(s) method with the Sedfit software.

In vitro ubiquitination assays. For the SdeA-mediated ubiquitination reaction, 0.1 μg His_6-SdeA and 1 μg GST–SidJ were preincubated in a 25-μl reaction system containing 50 mM Tris-HCl (pH 7.5), 1 mM DTT and 1 mM β-NAD^+ for 2 h at 37 °C. When needed, 5 mM $MgCl_2$, 1 mM L-glutamate, 1 mM ATP and 1 μM CaM (Sigma-Aldrich, C4874) were supplemented. After a 2-h preincubation, a cocktail containing 1 mM β-NAD^+, 0.25 μg 4×Flag–Rab33b and 5 μg ubiquitin was supplemented into the reactions and the reaction was allowed to proceed for another 2 h at 37 °C.

In vitro glutamylation assays. His_6-SdeA (0.1 μg) and GST–SidJ (1 μg) were incubated in a 25-μl reaction system containing 50 mM Tris-HCl (pH 7.5), 1 mM DTT, 5 mM $MgCl_2$, 1 mM L-glutamate, 1 mM ATP and 1 μM CaM for 2 h at 37 °C. To measure the glutamylase activity of SidJ using ^{14}C-glutamate, 2 μg His_6-SdeA and 0.5 μg GST–SidJ were incubated in a 25-μl reaction system containing 50 mM Tris-HCl (pH 7.5), 1 mM DTT, 5 mM $MgCl_2$, 1 μCi ^{14}C-L-glutamate (Perkin Elmer NEC290E050UC), 1 mM ATP and 1 μM CaM for 2 h at 37 °C. Products were resolved by SDS–PAGE and stained with Coomassie Brilliant Blue. Gels were then dried and signals were detected with X-ray films with a BioMax TranScreen LE (Kodak) for 3 days at −80 °C.

In vitro AMPylation assays. GST–SidJ (2 μg) was incubated in a 25-μl reaction system containing 50 mM Tris-HCl (pH 7.5), 1 μM CaM, 1 mM DTT, 5 mM $MgCl_2$ and 5 μCi ATP-α-^{32}P (Perkin Elmer BLU003H250UC) for 2 h at 37 °C. When needed, 3 μg His_6-SdeA and 1 mM L-glutamate were supplemented. Products were

resolved by SDS–PAGE and stained with Coomassie Brilliant Blue. Gels were then dried and signals were detected with X-ray films.

HPLC analysis of glutamylation reactions. SidJ(ΔN99) (40 μg) was incubated with 1 mM ATP in a 100-μl reaction system containing 50 mM Tris-HCl (pH 7.5) and 50 mM NaCl for 4 h at 37 °C. When needed, 1 mM CaM, 2 mM L-glutamate and 80 μg SdeA or SdeA(E860A) were supplemented. Samples were injected into a Waters Acquity UPLC system equipped with a C18 reversed-phase column and a UV detector. Components were eluted isocratically with 100% H_2O for 2 min followed by a 10-min gradient to 95% H_2O and 5% acetonitrile. ATP and AMP (1 mM) were run as standards.

Antibodies and immunoblotting. Purified His_6-GFP was used to raise rabbit-specific antibodies using a standard protocol (Pocono Rabbit Farm and Laboratory). The antibodies were affinity-purified as previously described[20]. Antibodies specific for SidJ and SdeA have been previously described[2,5]; commercial antibodies used are as follows: anti-Flag (Sigma-Aldrich, F1804; 1:2,000); anti-HA (Roche, 11867423001; 1:5,000), anti-ICDH[27] (1:10,000); anti-tubulin (DSHB, E7; 1:10,000); anti-HIF-1α (R&D Systems, MAB1536; 1:1,000); anti-PGK1 (Abcam, ab113687; 1:2,500); anti-CaM (Millipore, 05-173; 1:2,000). Membranes were then incubated with an appropriate IRDye infrared secondary antibody and scanned using an Odyssey infrared imaging system (Li-Cor's Biosciences).

Constitution of the SidJ–CaM complex and size-exclusion chromatography. Proteins purified as described above were further purified using a size-exclusion chromatography column (Superdex 200 increase 10/300; GE Healthcare) equilibrated with a washing buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl) on an AKTA pure system (GE Healthcare). To constitute the protein complex, purified SidJ and CaM were mixed at a 1:1.2 molar ratio at 4 °C for 1 h on a rotary shaker, and the complex was purified by size-exclusion chromatography using the above column. In each case, the proteins were eluted with the washing buffer. Fractions containing the protein of interest were pooled and used for further analysis.

Liquid chromatography–tandem mass spectrometry analysis. The Flag–mART domain was purified from HEK293T cells coexpressing GFP–SidJ or GFP. After separation by SDS–PAGE, gel slices containing the protein detected by silver staining were digested as described previously[35]. The digested peptides were analysed on a C18 reversed-phase column connected to a UPLC (Acquity, Waters) coupled to an Orbitrap mass spectrometer (Q-Exactive Plus, Thermo Fisher Scientific), using the same conditions as described previously[36]. Tandem mass spectra were converted to peak lists using DeconMSn[37] and submitted for blind posttranslational modification search using MODa[38] against the *L. pneumophila* sequences from GenBank. Post-translational-modification candidates were confirmed by manual inspection, looking for consistent mass shifts in b and y fragment series, and by reprocessing the data with MaxQuant[39] considering the specific modifications.

Microscale thermophoresis. The interaction between SidJ and CaM and the ATP-binding activity of SidJ were measured by microscale thermophoresis using the NanoTemper Monolith NT.115 instrument set at 20% LED and 20–40% IR-laser power. Laser on and off times were set at 30 s and 5 s, respectively. Each measurement consists of 16 reaction mixtures in which the concentration of fluorescent-labelled SidJ was set to be constant at 150 nM and the concentration of two fold-diluted CaM ranged from 20 μM to 0.61 nM. For ATP binding, the concentrations of ATP used were from 100 μM to 3.05 nM. The NanoTemper Analysis 2.2.4 software was used to fit the data and to determine the K_d.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The atomic coordinates and structure factors of the SidJ(Se-Met)–CaM, SidJ–CaM and SidJ–CaM–AMP have been deposited in the Protein Data Bank (PDB) under the accession codes 6K4L, 6K4K and 6K4R, respectively.

25. Berger, K. H. & Isberg, R. R. Two distinct defects in intracellular growth complemented by a single genetic locus in *Legionella pneumophila*. *Mol. Microbiol.* **7**, 7–19 (1993).
26. Luo, Z. Q. & Farrand, S. K. Signal-dependent DNA binding and functional domains of the quorum-sensing activator TraR as identified by repressor activity. *Proc. Natl Acad. Sci. USA* **96**, 9009–9014 (1999).
27. Xu, L. et al. Inhibition of host vacuolar H^+-ATPase activity by a *Legionella pneumophila* effector. *PLoS Pathog.* **6**, e1000822 (2010).
28. Sheedlo, M. J. et al. Structural basis of substrate recognition by a bacterial deubiquitinase important for dynamics of phagosome ubiquitination. *Proc. Natl Acad. Sci. USA* **112**, 15090–15095 (2015).
29. Mumberg, D., Müller, R. & Funk, M. Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* **156**, 119–122 (1995).
30. Otwinowski, Z. & Minor, W. [20] Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
31. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
32. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
33. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).
34. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
35. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V. & Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **1**, 2856–2860 (2006).
36. Gan, N., Nakayasu, E. S., Hollenbeck, P. J. & Luo, Z. Q. *Legionella pneumophila* inhibits immune signalling via MavC-mediated transglutaminase-induced ubiquitination of UBE2N. *Nat. Microbiol.* **4**, 134–143 (2019).
37. Mayampurath, A. M. et al. DeconMSn: a software tool for accurate parent ion monoisotopic mass determination for tandem mass spectra. *Bioinformatics* **24**, 1021–1023 (2008).
38. Na, S., Bandeira, N. & Paek, E. Fast multi-blind modification search through tandem mass spectrometry. *Mol. Cell. Proteomics* **11**, M111.010199 (2012).
39. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* **11**, 2301–2319 (2016).

Acknowledgements We thank C. Fan for assistance with structure determination and for discussions, and K. Weitz for assistance with the mass spectrometry analysis. This work was supported in part by National Institutes of Health grants R01AI127465 and R01GM126296, the National Natural Science Foundation of China grants 31770948, 31570875 and 31200559 (S.O.) and by a research fund from the First Hospital of Jilin University. Mass spectrometry analysis was performed in the Environmental Molecular Sciences Laboratory, a US Department of Energy national scientific user facility at Pacific Northwest National Laboratory. Battelle operates the Pacific Northwest National Laboratory for the Department of Energy under contract DE-AC05-76RL001830. The diffraction data were collected at beamline BL-17U1 of the Shanghai Synchrotron Radiation Facility.

Author contributions N.G. and Z.-Q.L. conceived the ideas for this work. Unless specified, N.G. and Yao Liu performed the experiments. Yao Liu, Yancheng Liu, N.G. and J.Q. performed the yeast experiments; G.M.F. and E.S.N. performed mass spectrometric analyses. X.Z., X.X., C.H., B.Z., L.Z. and S.O. determined the structures and analysed protein properties using biophysical tools. K.P. and C.D. performed HPLC analysis of nucleotide products. N.G., Yao Liu, E.S.N., S.O. and Z.-Q.L. interpreted the results. N.G., Yao Liu, S.O. and Z.-Q.L. wrote the manuscript and all authors provided editorial input.

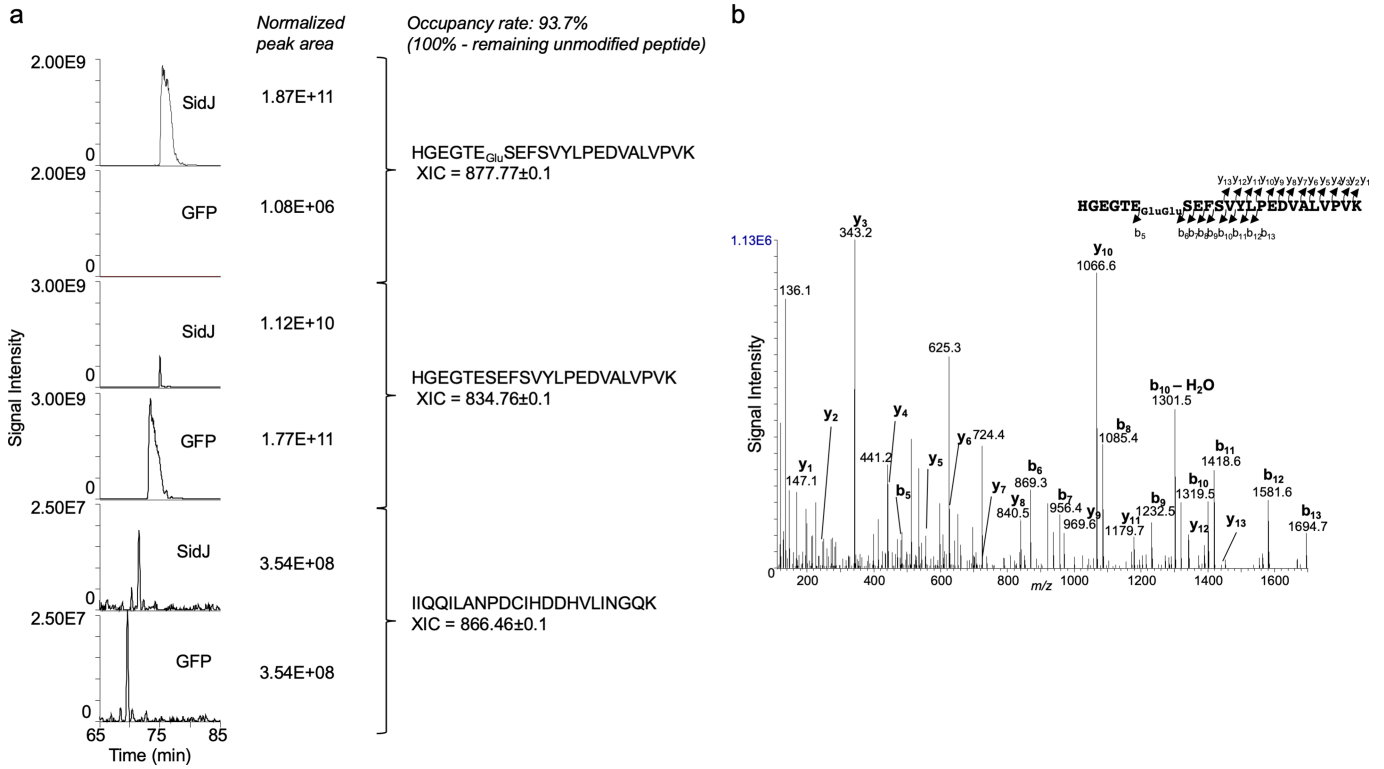
Competing interests The authors declare no competing interests.

Additional information

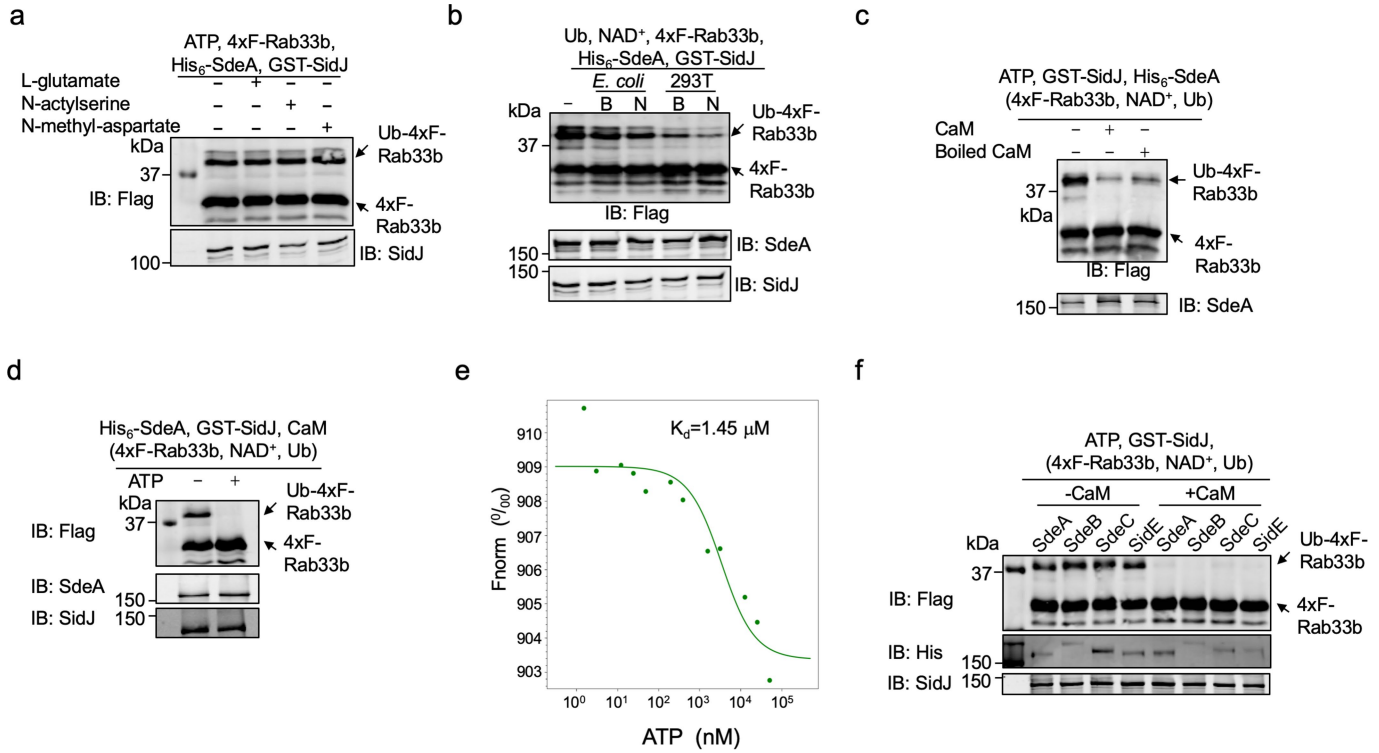
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1439-1.

Correspondence and requests for materials should be addressed to S.O. or Z.-Q.L.

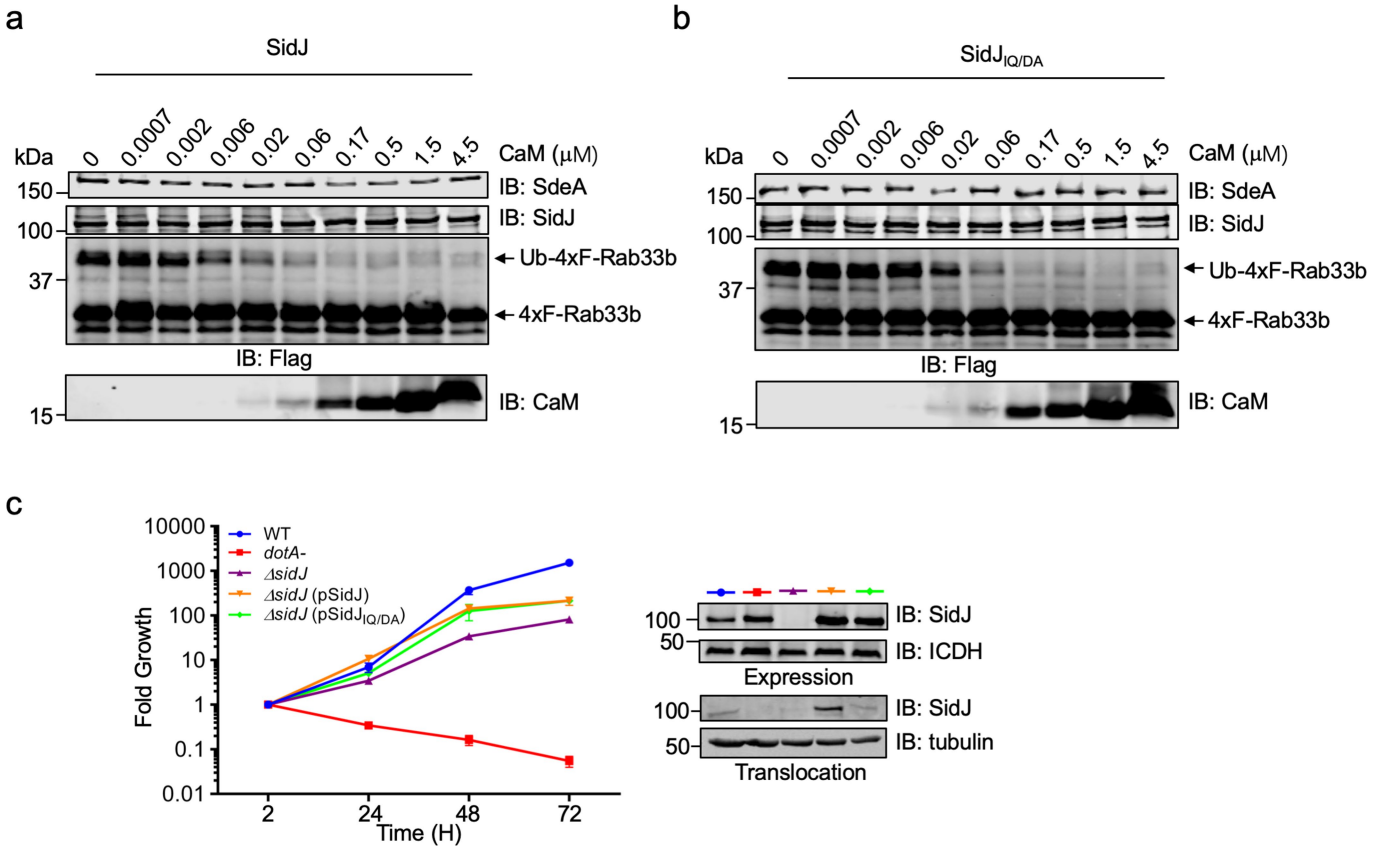
Peer review information *Nature* thanks Friedrich Förster, Elizabeth Hartland, Carsten Janke and the other anonymous reviewer(s) for their contribution to the peer review of this work.



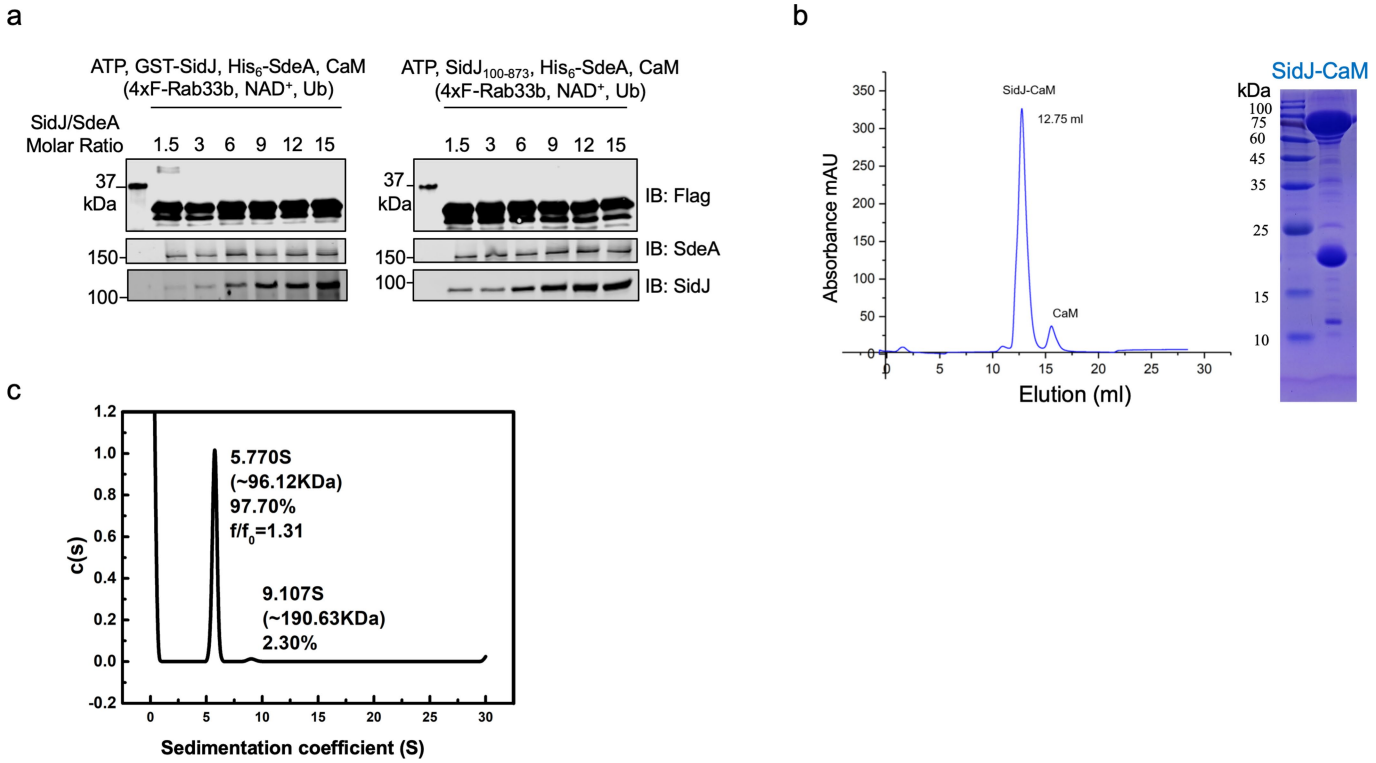
Extended Data Fig. 1 | Determination of the modification rate of E860 of SdeA. a, Peak areas of the extracted-ion chromatograms (XIC) were normalized on the basis of the area of the unmodified peptide –I_{608}IQQILANPDCIHDDHVLINGQK$_{630}$–. The occupancy rate of glutamylation on the residue was calculated on the basis of the consumption of the unmodified –H_{855}GEGTESEFSVYLPEDVALVPVK$_{877}$– in samples from cells cotransfected to express GFP–SidJ compared to those of controls from cells transfected to express GFP. **b**, SidJ induces a 258.09-Da post-translation modification on E860 within the mART motif of SdeA. 4×Flag–mART purified from HEK293T cells coexpressing SidJ detected by silver staining (Fig. 2d) was analysed by mass spectrometric analysis. The tandem mass spectrum shows the fragmentation profile of the modified peptide –H_{855}GEGTE$_{GluGlu}$SEFSVYLPEDVALVPVK$_{877}$–, including ions b_5 and b_6, which confirms the modification site at the E860 residue. In each case, similar results were obtained in three independent experiments.



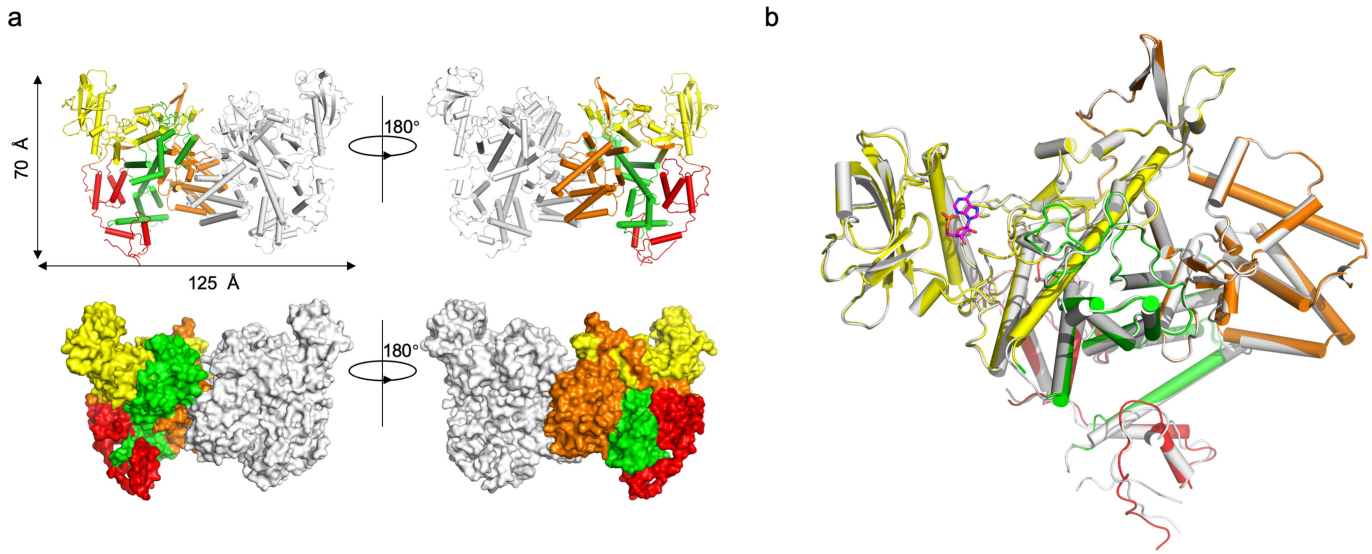
Extended Data Fig. 2 | The effects of cell lysates, ATP and heat treatment of CaM on the activity of SidJ and its inhibition of the activity of all members of the SidE family. a, Inhibition of SdeA activity does not occur in in vitro reactions containing L-glutamate or each of its two structural isomers. L-glutamate, *N*-acetylserine or *N*-methylaspartate was incubated with SdeA, SidJ and ATP for 2 h before assaying for the activity of SdeA. **b**, One or more factors from mammalian cells are required for SidJ to inhibit SdeA. Lysates from *E. coli* or HEK293T cells were added to reactions containing SdeA and SidJ for 2 h before measuring the activity of SdeA. **c**, Heat treatment does not completely abolish CaM activity. CaM or CaM treated by heating at 100 °C for 5 min was included in reactions that allow glutamylation of SdeA for 2 h. A cocktail containing 4×Flag–Rab33b, NAD^+ and ubiquitin was added to each reaction. Samples were resolved by SDS–PAGE and analysed for Rab33b ubiquitination after another 2 h incubation at 37 °C. **d**, The activity of SidJ requires ATP. His_6-SdeA was incubated with GST–SidJ, L-glutamate and CaM in reactions with or without 1 mM ATP for 2 h; 4×Flag–Rab33b, NAD^+ and ubiquitin were added to each reaction. After another 2-h incubation, the activity of SdeA was evaluated by the production of ubiquitinated Ra33b. Protein components in the reactions were detected by immunoblotting with specific antibodies. **e**, The binding of ATP by SidJ. Binding of ATP by purified SidJ was evaluated using microscale thermophoresis in which the concentration of SidJ was kept constant. K_d was determined by the NanoTemper Analysis 2.2.4 software. **f**, SidJ inhibits the activity of members of the SidE family. A recombinant protein of each member of the SidE family was incubated with ATP, L-glutamate and GST–SidJ in the presence or absence of CaM for 2 h, and a cocktail containing 4×Flag–Rab33b, NAD^+ and ubiquitin was added to the reactions. After an additional 2-h incubation, modification of Rab33b was detected by immunoblotting with a Flag-specific antibody. The formation of Ub-4×Flag–Rab33b is indicated by a shift in molecular mass. In each panel, data shown are one representative from at least three independent experiments with similar results.



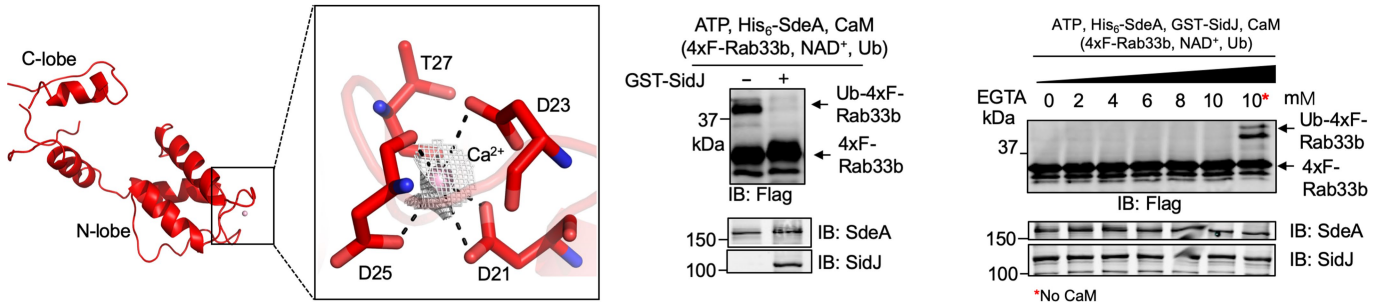
Extended Data Fig. 3 | The IQ motif of SidJ is required for its optimal response to CaM. a, **b**, The IQ motif is required for the optimal activity of SidJ in response to CaM. Serially diluted CaM was preincubated with SidJ (**a**) or the SidJ(I841D/Q842A) mutant (**b**) and SdeA in the glutamylation buffer at 37 °C for 2 h. A cocktail containing 4×Flag–Rab33b, NAD^+ and ubiquitin was added to the reactions. After incubation for another 2 h at 37 °C, proteins separated by SDS–PAGE were assessed using the indicated antibodies. In each panel, data shown are one representative from at least three independent experiments with similar results. **c**, The SidJ(I841D/Q842A) mutant complements the intracellular growth defect of the *ΔsidJ* mutant. *A. castellanii* was infected with the indicated bacterial strains and intracellular bacteria were determined at the indicated time points. Experiments on each strain were performed in triplicate and similar results were obtained in two independent experiments. Results are from one representative experiment performed in triplicate from three independent experiments; error bars represent s.e.m. ($n = 3$).



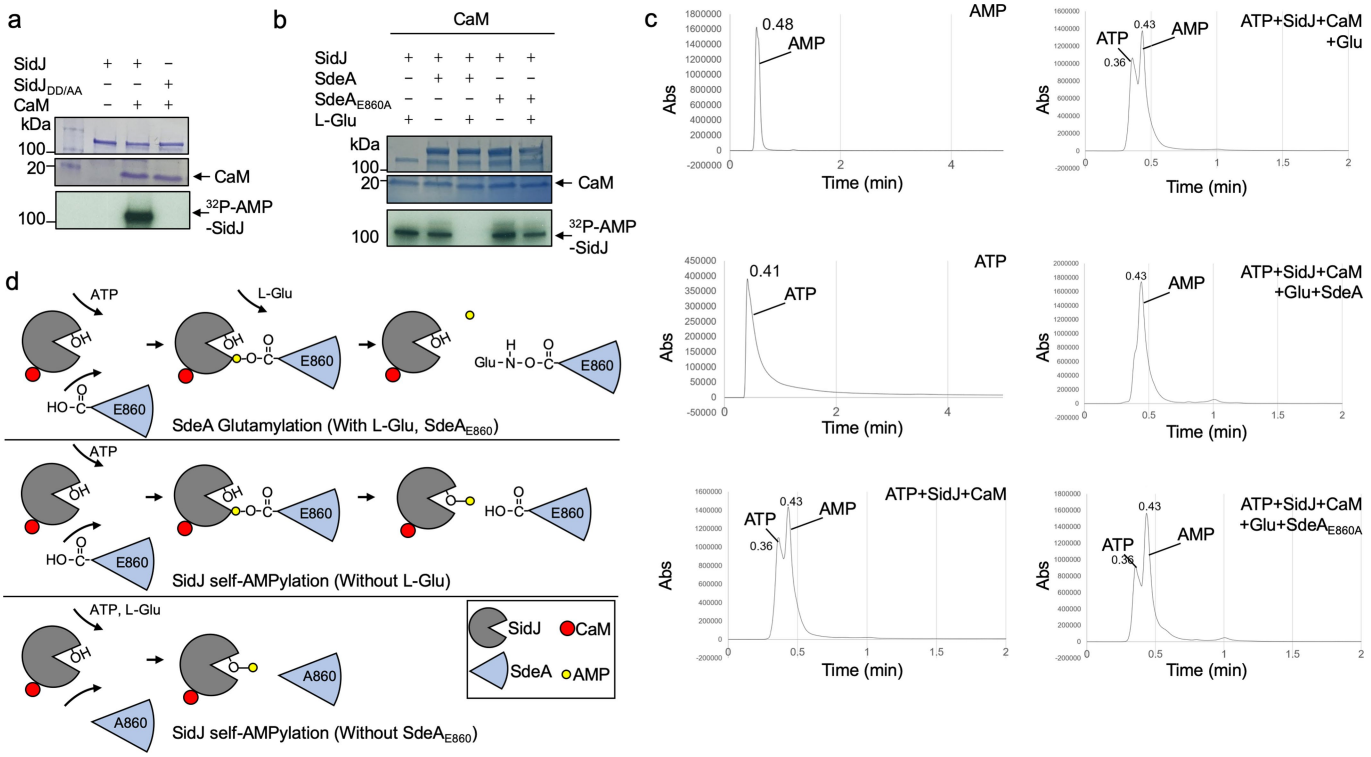
Extended Data Fig. 4 | SidJ forms a stable heterodimer with CaM at a molar ratio of 1:1. a, SidJ(ΔN99) maintains the ability to inhibit SdeA activity, to a similar extent to that of full-length SidJ. SdeA was incubated with GST–SidJ or SidJ(ΔN99) at indicated molar ratios in reactions containing ATP, L-glutamate, CaM for 2 h at 37 °C. A cocktail containing 4×Flag–Rab33b, NAD^+ and ubiquitin was added to each reaction for an additional 2 h at 37 °C, and the proteins resolved by SDS–PAGE were analysed with the indicated antibodies. SdeA activity was measured by the production of ubiquitinated Rab33b as indicated by a shift in molecular mass. **b**, Size-exclusion chromatography profiles of SidJ–CaM. Left, purified proteins were separated by a Superdex 200 Increase 10/300 column (GE Healthcare) on an AKTA pure system. Right, fractions with strong absorbance at an optical density of 260 nm (OD_{260}) were collected and analysed by SDS–PAGE followed by detection with Coomassie brilliant blue staining. **c**, The heterodimer formed between SidJ(ΔN99) with CaM is a monomer. Analytical ultracentrifugation analysis yielded a sedimentation coefficient of 5.770 S, and a molecular mass of approximately 96.12 kDa, which is indicative of the heterodimer of SidJ(ΔN99) and CaM. In each panel, data shown are one representative from at least three independent experiments with similar results.



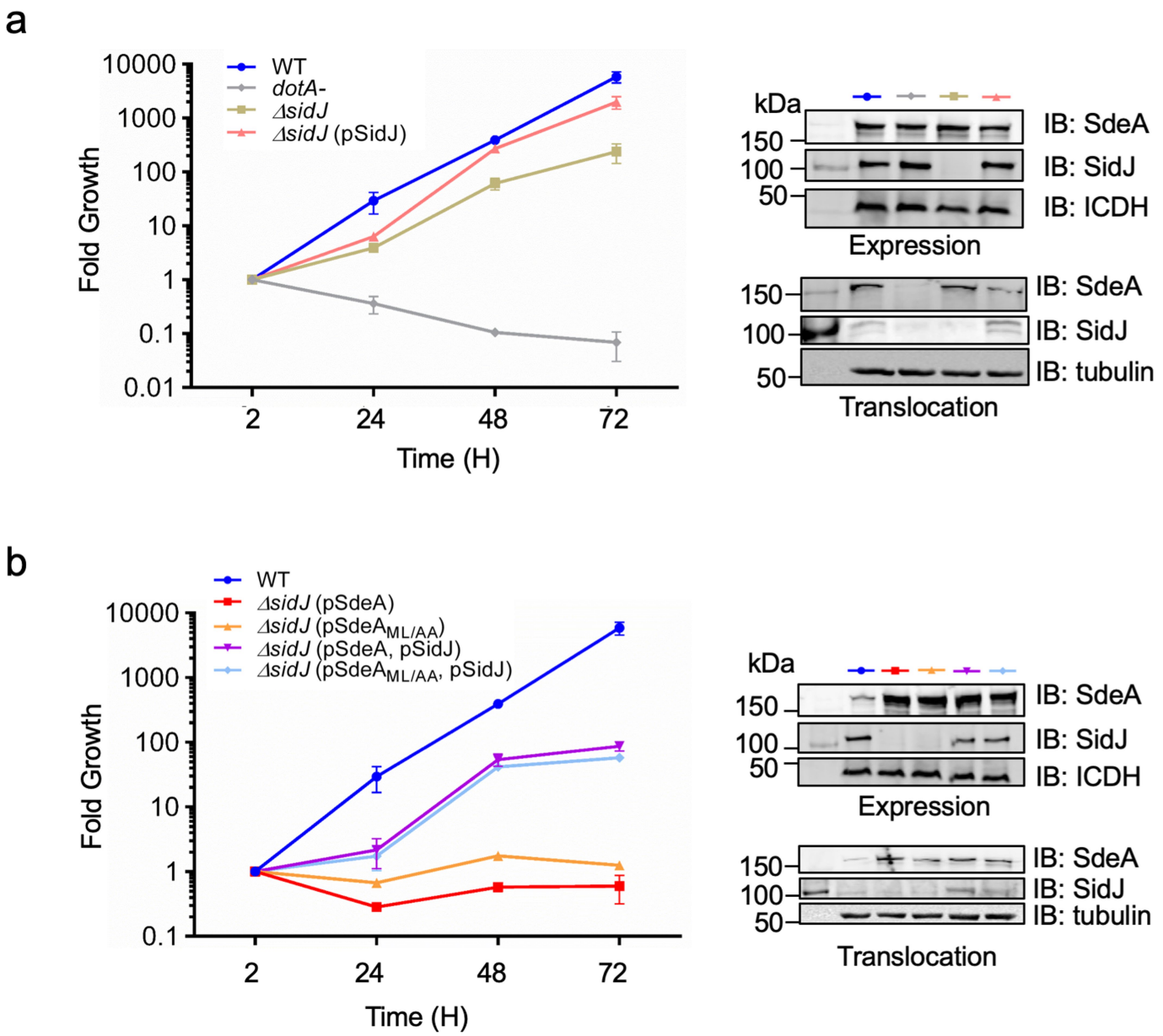
Extended Data Fig. 5 | Overall structure of the SidJ–CaM complex in one asymmetric unit and the comparison of complex structures with or without AMP. a, Two views of the structure of the SidJ–CaM heterodimer in the asymmetric unit displayed as a ribbon diagram (top) and with surface rendering (bottom); one of the SidJ–CaM heterodimers is coloured as shown in Fig. 4 and the other one is coloured in grey. **b**, Superimposition of the structures of SidJ–CaM and SidJ–CaM–AMP. The SidJ–CaM–AMP ternary complex is coloured as shown in Fig. 4d and the SidJ–CaM binary complex is coloured in grey.



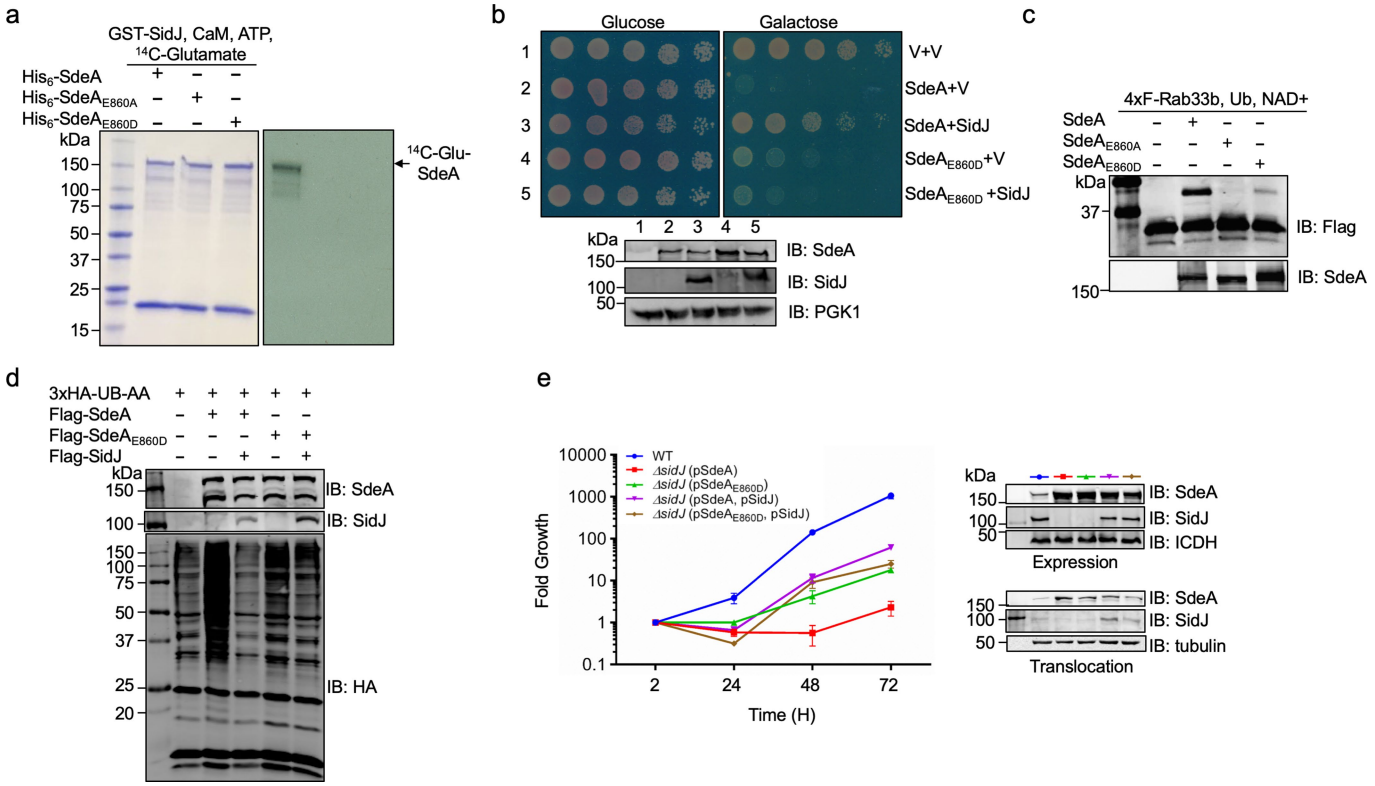
Extended Data Fig. 6 | Interactions between CaM and Ca^{2+} from the crystal structures and the role of Ca^{2+} in the activation of SidJ by CaM. **a**, Key residues of CaM involved in the interaction with Ca^{2+}. Ca^{2+} is coordinated by D21, D23, D25 and T27 of CaM, which are shown as red sticks. Ca^{2+} is shown as a pink sphere. Electron density of a simulated annealing $F_o - F_c$ omit map for Ca^{2+} contoured at 3.0σ. **b**, Dialysis against 20 mM EGTA does not abolish the activity of SidJ. All proteins used in the reactions were dialysed against a buffer containing 20 mM EGTA for 14 h. SdeA was incubated with SidJ in reactions containing ATP and EGTA-dialysed CaM for 2 h at 37 °C. Reactions without SidJ were established as a control. A cocktail containing 4×Flag–Rab33b, NAD^+ and ubiquitin was added to each reaction. After further incubation for 2 h at 37 °C, proteins resolved by SDS–PAGE were analysed with the indicated antibodies. SdeA activity was measured by the production of ubiquitinated Rab33b as indicated by a shift in molecular mass. **c**, The activity of SidJ is not sensitive to 10 mM EGTA. SdeA was first incubated with SidJ for glutamylation with the indicated amounts of EGTA for 2 h at 37 °C. NAD^+, 4×Flag–Rab33b and ubiquitin were then supplemented to the reactions, which were allowed to proceed for 2 h at 37 °C before resolution by SDS–PAGE. Rab33b modification was detected as described in **b**. Proteins in the reactions were detected by immunoblotting with specific antibodies. In **b**, **c**, similar results were obtained in at least three independent experiments.



Extended Data Fig. 7 | The mechanism of SidJ-induced CaM-dependent self-AMPlyation and SdeA glutamylation. a, SidJ induces self-AMPylation in a CaM-dependent manner. SidJ was incubated with ^{32}P-α-ATP and Mg^{2+}, with or without CaM for 2 h at 37 °C. After separation by SDS–PAGE, the incorporation of ^{32}P-α-ATP was detected by autoradiography. **b**, SdeA glutamylation by SidJ interferes with SidJ self-AMPylation. SidJ was incubated with ^{32}P-α-ATP, Mg^{2+} and CaM for 2 h at 37 °C. L-glutamate, SdeA and SdeA(E860D) were supplemented as stated. After separation by SDS–PAGE, the incorporation of ^{32}P-α-ATP was detected by autoradiography. **c**, SdeA glutamylation by SidJ accelerates ATP hydrolysis and the release of AMP. SidJ was incubated with the indicated components for 2 h at 37 °C. Samples were analysed by HPLC. AMP and ATP were used as standards. In **a**–**c**, data shown are one representative from at least three independent experiments with similar results. **d**, Schematic model of SidJ-induced glutamylation and AMPylation. SidJ induces glutamylation on SdeA(E860D) when ATP and L-glutamate are supplemented into the reaction. In reactions in which L-glutamate or modifiable SdeA are not present, SidJ undergoes self-AMPylation.



Extended Data Fig. 8 | Intracellular growth phenotypes associated with the Δ*sidJ* mutant expressing SdeA and its mutants. a, Intracellular defects of the *L. pneumophila* Δ*sidJ* mutant can be complemented by SidJ expressed from a multicopy plasmid. The indicated strains were used to infect *A. castellanii* at an MOI of 0.05 and the growth of the bacteria was evaluated at 24-h intervals. Fold growth was calculated on the basis of total bacterial counts at the indicated time points and those of the 2-h time point. **b**, Overexpression of a SdeA mutant defective in substrate recognition inhibits intracellular growth of the Δ*sidJ* mutant. Intracellular growth of the indicated *L. pneumophila* strains in *A. castellanii* was evaluated as described in **a**. In each panel, the expression of SidJ, SdeA and its mutants in bacterial cells and their translocation into infected cells was determined by immunoblotting from total bacterial cell lysates and the saponin-soluble fraction of infected cells, with isocitrate dehydrogenase and tubulin as loading controls, respectively (right). In each case, results are from one representative experiment performed in triplicate from three independent experiments; error bars represent s.e.m. ($n = 3$).



Extended Data Fig. 9 | SidJ functions to regulate the activity of SdeA during *L. pneumophila* infection. a, SdeA(E860D) is resistant to glutamylation catalysed by SidJ. SdeA, SdeA(E860A) or SdeA(E860D) was added to reactions containing GST–SidJ, ^{14}C-glutamate ATP and CaM and the reactions were allowed to proceed for 2 h at 37 °C. After separation by SDS–PAGE, the incorporation of ^{14}C-glutamate was detected by autoradiography. **b**, Yeast toxicity induced by SdeA(E860D) cannot be suppressed by SidJ. A plasmid that directs the expression of SidJ was introduced into yeast strains expressing SdeA or SdeA(E860D) from a galactose inducible promoter, serially diluted yeast cells were spotted onto glucose or galactose medium for 2 days and the growth of the cells was evaluated by imaging (top). The expression of SidJ, SdeA and SdeA(E860D) was determined by immunoblotting with specific antibodies. The PGK1 (3-phosphoglyceric phosphokinase-1) was analysed as a loading control (bottom). **c**, SdeA(E860D) still ubiquitinates Rab33b. Reactions containing the indicated components were allowed to proceed for 2 h at 37 °C, samples were then resolved by SDS–PAGE and ubiquitination of Rab33b was assessed by immunoblotting with a Flag-specific antibody to detect the production of modified Rab33b with a higher molecular mass. **d**, SdeA(E860D)-mediated protein ubiquitination in mammalian cells is insensitive to SidJ. HEK293T cells were transfected to express the indicated proteins for 16–18 h. Cleared cell lysates were subjected to SDS–PAGE and immunoblotting with an HA-specific antibody to detect proteins ubiquitinated by 3×HA–Ub-AA. The amounts of SdeA, SdeA(E860D) and SidJ were assessed by antibodies specific for these proteins. Note that coexpression of SidJ reduced the ubiquitination induced by SdeA but not by SdeA(E860D). In **a**–**d**, data shown are one representative from at least three independent experiments with similar results. **e**, The effects of SidJ on intracellular growth defect caused by overexpression of SdeA or SdeA(E860D). The indicated *L. pneumophila* strains were used to infect *A. castellanii* at an MOI of 0.05 and the growth of the bacteria was evaluated at 24-h intervals. Fold growth was calculated on the basis of total bacterial counts at the indicated time points. Note the difference between strain Δ*sidJ* (pSdeA) and Δ*sidJ* (pSdeA, pSidJ). The growth defect caused by overexpressing the SdeA(E860D) mutant cannot be rescued by SidJ. The amounts of relevant proteins in bacterial cells and in infected cells were analysed by immunoblotting from total bacterial cell lysates and the saponin-soluble fraction of infected cells, with isocitrate dehydrogenase and tubulin as loading controls, respectively (right). Results showen are from one representative experiment performed in triplicate from three independent experiments; error bars represent s.e.m. ($n = 3$).

Extended Data Table 1 | Data collection and refinement statistics

	$SidJ_{Se\text{-}Met}$-CaM (PDB 6K4L)	SidJ-CaM (PDB 6K4K)	SidJ-CaM-AMP (PDB 6K4R)
Data Collection			
Space group	$P\ 1\ 2_1\ 1$	$P\ 1\ 2_1\ 1$	$P\ 1\ 2_1\ 1$
Cell dimensions			
a, *b*, *c* (Å)	61.06, 159.25, 135.81	60.96, 159.53, 135.61	60.85, 159.18, 135.03
α, β, γ (°)	90.00, 101.68, 90.00	90.00, 101.89, 90.00	90.00, 101.78, 90.00
Wavelength (Å)	0.9792	0.9792	0.9792
Resolution (Å)	66.50-2.95 (3.01-2.95) *	55.46-2.71 (2.81-2.71)	66.09-3.11 (3.22-3.11)
R_{merge}	0.158 (0.959)	0.176 (1.401)	0.230 (1.131)
$I / \sigma I$	12.8 (2.5)	12.1 (2.2)	12.5 (2.3)
Completeness (%)	99.90 (100.00)	96.87 (97.82)	92.45 (99.80)
Redundancy	6.8 (7.1)	12.9 (12.2)	6.5 (5.4)
Refinement			
Resolution (Å)	2.95	2.71	3.11
No. reflections	53475	66262	41852
R_{work} / R_{free}	0.252/0.278	0.205/0.243	0.239/0.279
No. atoms			
Protein	12936	12738	12640
Ligand/ion	2	2	100
Water	16	2	0
B factors (Å^2)			
Protein	58.30	69.39	65.00
Ligand/ion	64.10	128.39	75.65
R.m.s. deviations			
Bond lengths (Å)	0.005	0.006	0.003
Bond angles (°)	0.89	0.93	0.63

*For each structure one crystal was used. Values in parentheses are for highest-resolution shell.

nature research

Corresponding author(s): Dr. Zhao-Qing Luo

Last updated by author(s): Jul 7, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Odyssey Licor imaging system, GE Healthcare AKTA, Orbitrap mass spectrometer (Q-Exactive Plus, Thermo Fisher Scientific), NanoTemper Monolith NT.115, Beamline BL-17U1 of the Shanghai Synchrotron Radiation Facility (SSRF), Beckman XL-A analytical ultracentrifuge equipped with absorbance optics and an An60 Ti rotor, Waters Acquity UPLC equipped with a C18 reversed-phase column and a UV detector
Data analysis	Microsoft Excel 2016, and Prism Graphpad 7.0, GE Healthcare Unicorn V7.3, HKL2000 software, AutoSol implemented in PHENIX V 1.5, PROCHECK V 3.5.4, PDBePISA v1.52, Coot V 0.8.9, Sedfit Software V16.1c, Odyssey Image Studio V 5.2, DeconMSn V2.2, MODa V1.03, MaxQuant v1.6.6.0, NanoTemper Analysis 2.2.4 software, Pymol V1.8.6.2

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data that support the conclusions of this study are included in this published article along with its Supplementary Information files, and are also available from the corresponding author upon request. The atomic coordinates and structure factors of the SidJSe-Met-CaM, SidJ-CaM and SidJ-CaM-AMP have been deposited in the Protein Data Bank (PDB) under the accession codes 6K4L, 6K4K and 6K4R, respectively.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size is 3
Data exclusions	No data were excluded
Replication	For growth curve experiments, infections with each bacterial strain was performed in triplicate and similar results were obtained in two independent experiments. All other experiments were performed for at least 3 times.
Randomization	Randomisation was not required as no human participants or animal models were reported in this manuscript. The experiments were performed on matched cell lines or specific biochemical reactions.
Blinding	Blinding was not used as the present study is not a clinical research trial

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Purified His6-GFP was used to raise rabbit specific antibodies using a standard protocol (Pocono Rabbit Farm & Laboratory). The antibodies were affinity purified as describe Antibodies specific for SidJ and SdeA had been described. Commercial antibodies used are listed as below: anti-Flag (Sigma, Cat# F1804), 1: 2000; anti-HA (Roche, cat# 11867423001 1:5,000), anti-ICDH3, 1:10,000, anti-tubulin (DSHB, E7) 1:10,000, anti-HIF-1a (R&D systems, cat#MAB1536 1:1,000), anti-PGK1 (Abcam, cat# ab113687 1:2,500), anti-CaM (Millipore, cat#05-173 1:2,000). Membranes were then incubated with an appropriate IRDye infrared secondary antibody (Invitrogen Goat anti-Mouse IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 680 CAT#A21057 1:10,000; Invitrogen Goat anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 680, cat#A21109, 1:10,000; LICOR IRDye® 800CW Goat anti-Rabbit IgG (H + L), CAT#926-32211 1:10,000; LICOR IRDye® 800CW Goat anti-Mouse IgG (H + L), CAT#926-32210 1:10,000) and scanned using an Odyssey infrared imaging system (Li-Cor's Biosciences).
Validation	GFP and ICDH antibodies were described in: Xu, L. et al. Inhibition of host vacuolar H+-ATPase activity by a Legionella pneumophila effector. PLoS Pathog. 6, e1000822 (2010). SdeA antibody was described in: Qiu, J. et al. Ubiquitination independent of E1 and E2 enzymes by bacterial effectors. Nature 533, 120-124, doi:10.1038/nature17657 (2016). SidJ antibody was described in: Liu, Y. & Luo, Z. Q. The Legionella pneumophila effector SidJ is required for efficient recruitment of endoplasmic reticulum proteins to the bacterial phagosome. Infect Immun 75, 592-603 (2007). Antibody, catalogue number, manufacturer information for commercial antibodies: anti-Flag (Sigma, Cat# F1804): https://www.sigmaaldrich.com/catalog/product/sigma/f1804?lang=en&region=US anti-HA (Roche, cat# 11867423001): https://www.sigmaaldrich.com/catalog/product/roche/roahaha?lang=en&region=US anti-tubulin (DSHB, E7): http://dshb.biology.uiowa.edu/tubulin-beta-_2

anti-HIF-1a (R&D systems, cat#MAB1536): https://www.rndsystems.com/products/human-mouse-rat-hif-1alpha-antibody-241809_mab1536
anti-PGK1 (Abcam, cat# ab113687): https://www.abcam.com/pgk1-antibody-22c5d8-ab113687.html
anti-CaM (Millipore, cat#05-173): http://www.emdmillipore.com/US/en/product/Anti-Calmodulin-Antibody,MM_NF-05-173
Invitrogen Goat anti-Mouse IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 680 CAT#A21057: https://www.thermofisher.com/antibody/product/Goat-anti-Mouse-IgG-H-L-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-21057
Invitrogen Goat anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 680, cat#A21109: https://www.thermofisher.com/antibody/product/Goat-anti-Rabbit-IgG-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-21109
LICOR IRDye® 800CW Goat anti-Rabbit IgG (H + L), CAT#926-32211: https://www.licor.com/bio/reagents/irdye-800cw-goat-anti-rabbit-igg-secondary-antibody
LICOR IRDye® 800CW Goat anti-Mouse IgG (H + L), CAT#926-32210: https://www.licor.com/bio/reagents/irdye-800cw-goat-anti-mouse-igg-secondary-antibody

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	HEK293T cells, Acanthamoeba castellanii cells, Yesat W303 and BY4741 cells were purchased from ATCC
Authentication	Authenticated by ATCC. ATCC uses morphology, karyotyping, and PCR based approaches to confirm the identity of human cell lines and to rule out both intra- and interspecies contamination. These include an assay to detect species specific variants of the cytochrome C oxidase I gene (COI analysis) to rule out inter-species contamination and short tandem repeat (STR) profiling to distinguish between individual human cell lines and rule out intra-species contamination.
Mycoplasma contamination	Contamination tested by using the universal mycoplasma detection kit from ATCC (cat# 30-1012K). All cell lines tested are confirmed as negative for mycoplasma contamination
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used

LETTER

https://doi.org/10.1038/s41586-019-1456-0

CD24 signalling through macrophage Siglec-10 is a target for cancer immunotherapy

Amira A. Barkal[1,2,3,4], Rachel E. Brewer[1,2,3], Maxim Markovic[1,2,3], Mark Kowarsky[5], Sammy A. Barkal[1], Balyn W. Zaro[1,2,3], Venkatesh Krishnan[6], Jason Hatakeyama[1,7], Oliver Dorigo[6], Layla J. Barkal[8] & Irving L. Weissman[1,2,3,9]*

Ovarian cancer and triple-negative breast cancer are among the most lethal diseases affecting women, with few targeted therapies and high rates of metastasis. Cancer cells are capable of evading clearance by macrophages through the overexpression of anti-phagocytic surface proteins called 'don't eat me' signals—including CD47[1], programmed cell death ligand 1 (PD-L1)[2] and the beta-2 microglobulin subunit of the major histocompatibility class I complex (B2M)[3]. Monoclonal antibodies that antagonize the interaction of 'don't eat me' signals with their macrophage-expressed receptors have demonstrated therapeutic potential in several cancers[4,5]. However, variability in the magnitude and durability of the response to these agents has suggested the presence of additional, as yet unknown 'don't eat me' signals. Here we show that CD24 can be the dominant innate immune checkpoint in ovarian cancer and breast cancer, and is a promising target for cancer immunotherapy. We demonstrate a role for tumour-expressed CD24 in promoting immune evasion through its interaction with the inhibitory receptor sialic-acid-binding Ig-like lectin 10 (Siglec-10), which is expressed by tumour-associated macrophages. We find that many tumours overexpress CD24 and that tumour-associated macrophages express high levels of Siglec-10. Genetic ablation of either CD24 or Siglec-10, as well as blockade of the CD24–Siglec-10 interaction using monoclonal antibodies, robustly augment the phagocytosis of all CD24-expressing human tumours that we tested. Genetic ablation and therapeutic blockade of CD24 resulted in a macrophage-dependent reduction of tumour growth in vivo and an increase in survival time. These data reveal CD24 as a highly expressed, anti-phagocytic signal in several cancers and demonstrate the therapeutic potential for CD24 blockade in cancer immunotherapy.

CD24, also known as heat stable antigen or small-cell lung carcinoma cluster 4 antigen, is a heavily glycosylated glycosylphosphatidylinositol-anchored surface protein[6,7]. It is known to interact with Siglec-10 on innate immune cells to dampen damaging inflammatory responses to infection[8], sepsis[9], liver damage[10] and chronic graft versus host disease[11]. The binding of CD24 to Siglec-10 elicits an inhibitory signalling cascade, which is mediated by Src homology region 2 domain-containing phosphatases, SHP-1 and/or SHP-2. These phosphatases are associated with the two immunoreceptor tyrosine-based inhibition motifs in the cytoplasmic tail of Siglec-10, thereby blocking Toll-like-receptor-mediated inflammation and the cytoskeletal rearrangement required for cellular engulfment by macrophages[12–14]. Studies have shown that CD24 is expressed by several solid tumours[15,16]; however, a role for CD24 in modulating tumour immune responses has not yet been shown. We therefore sought to investigate whether CD24-mediated inhibition of the innate immune system could be harnessed by cancer cells as a mechanism of avoiding clearance by macrophages that express Siglec-10.

To assess the role of CD24–Siglec-10 signalling in regulating the macrophage-mediated immune response to cancer, we examined the expression of CD24 and Siglec-10 in various tumours and associated immune cells. RNA-sequencing data from The Cancer Genome Atlas (TCGA) and the Therapeutically Applicable Research to Generate Effective Treatment Program (TARGET) revealed high expression of *CD24* in nearly all tumours analysed (Extended Data Fig. 1a), as well as broad upregulation of *CD24* expression in several tumours as compared to known innate immune checkpoints (Fig. 1a). The largest upregulation of *CD24*—a $\log_2$ fold increase of more than nine—was observed in ovarian cancer; in addition, *CD24* expression was significantly higher in triple-negative breast cancer (TNBC) than in healthy breast cells or in oestrogen- and progesterone-receptor-positive (ER^+PR^+) breast cancers (Extended Data Fig. 1b, c). Stratification of patients by *CD24* expression revealed increased relapse-free survival for patients with ovarian cancer and an overall survival advantage for patients with breast cancer with lower *CD24* expression (Fig. 1b, c). We investigated *CD24* and *SIGLEC10* expression at a cellular level within the tumour by using single-cell RNA-sequencing data from six primary samples of TNBC[17] (NCBI Sequence Read Archive: PRJNA485423; Fig. 1d, Extended Data Fig. 1d, e). TNBC cells exhibited robust expression of *CD24*, whereas its expression was weak in all other cell clusters, thus illustrating the potential of CD24 as a tumour-specific marker (Fig. 1d). A substantial fraction of tumour-associated macrophages (TAMs) were found to express *SIGLEC10*, indicating the possibility of CD24–Siglec-10 interactions in TNBC (Fig. 1d). *CD24* expression was substantially higher than *PD-L1* (also known as *CD274*) expression in all patients analysed (Extended Data Fig. 1f), whereas *CD47* was highly expressed by all cell types (Fig. 1d). Fluorescence-activated cell sorting (FACS) analyses of primary human tumours revealed robust expression of the CD24 protein in breast cancer cells and ovarian cancer cells, and TAMs from both tumour types were found to express Siglec-10 (Fig. 1e, f, Extended Data Fig. 2a). Human peritoneal macrophages obtained from patients without cancer expressed low levels of Siglec-10 (Extended Data Fig. 2b). Analysis of subsets of peripheral blood mononuclear cells revealed low expression of Siglec-10 and CD24 in T cells, natural killer cells and monocytes, whereas B cells were found to express modest levels of Siglec-10 and high levels of CD24 (Extended Data Fig. 2c, d).

To investigate a role for CD24–Siglec-10 signalling in regulating the macrophage-mediated anti-tumour immune response (Fig. 2a), we engineered a polyclonal subline of the normally CD24-positive MCF-7 human breast cancer cell line that was deficient in CD24 (ΔCD24). Although unstimulated (M0) human donor-derived macrophages expressed low levels of Siglec-10 as measured by FACS, the addition of two inhibitory cytokines—TGFβ1 and IL-10—induced robust expression of Siglec-10, indicating that Siglec-10 expression may be regulated by TAM-specific gene-expression programs[18] (Extended Data Fig. 2e). Macrophages stimulated by TGFβ1 and IL-10 (M2-like)

[1]Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, CA, USA. [2]Ludwig Center for Cancer Stem Cell Research and Medicine, Stanford University School of Medicine, Stanford, CA, USA. [3]Stanford Cancer Institute, Stanford University School of Medicine, Stanford, CA, USA. [4]Stanford Medical Scientist Training Program, Stanford University, Stanford, CA, USA. [5]Department of Physics, Stanford University, Stanford, CA, USA. [6]Department of Obstetrics and Gynecology, Division of Gynecologic Oncology, Stanford University School of Medicine, Stanford, CA, USA. [7]Department of Urology, Stanford University School of Medicine, Stanford, CA, USA. [8]Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA. [9]Department of Pathology, Stanford University School of Medicine, Stanford, CA, USA. *e-mail: irv@stanford.edu

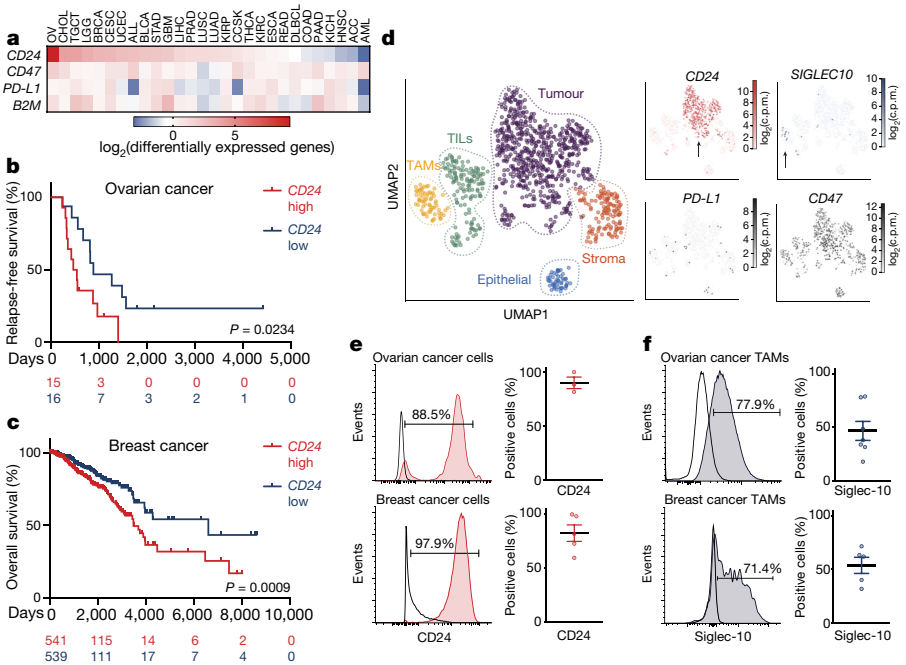


Fig. 1 | CD24 is overexpressed by human cancers and is co-expressed with Siglec-10 on TAMs. a, Heat map of *CD24* tumour to matched normal expression ratios ($\log_2$(differentially expressed genes)) compared to known immune checkpoints. Tumour study abbreviations and *n* values are provided in Supplementary Table 1. **b**, **c**, Relapse-free survival of patients with ovarian cancer ($n = 31$) (**b**) and overall survival of patients with breast cancer ($n = 1{,}080$) (**c**) with high or low *CD24* expression as defined by the median. Two-sided *P* value computed by a log-rank (Mantel–Cox) test. Numbers of subjects at risk in the high group (red) compared with the low group (blue) are indicated below the *x* axes. **d**, Uniform manifold approximation and projection (UMAP) dimension 1 and 2 plots displaying TNBC cells from 6 patients ($n = 1{,}001$ single cells). Left, cells are coloured by cluster identity; right, *CD24* (red) and *SIGLEC10* (blue) expression overlaid onto UMAP space as compared to the expression of *CD47* (grey) and *PD-L1* (grey). **e**, Left, representative histogram (obtained from flow cytometry results) of CD24 expression by ovarian cancer cells (top) or breast cancer cells (bottom); right, frequency of $CD24^+$ cancer cells in ovarian cancer ($n = 3$ donors) (top) or breast cancer ($n = 5$ donors) (bottom). Data are mean ± s.e.m. **f**, Left, representative histogram measuring the expression of Siglec-10 by ovarian cancer TAMs (top) or breast cancer TAMs (bottom); right, frequency of Siglec-10^+ TAMs in ovarian cancer ($n = 6$ donors) (top) or breast cancer ($n = 5$ donors) (bottom). Data are mean ± s.e.m.

were less phagocytic than unstimulated macrophages at baseline levels (Extended Data Fig. 2f). We found that stimulation with the classic M2-polarizing cytokine IL-4 was also sufficient to induce Siglec-10 expression (Extended Data Fig. 2g). Co-culture of either wild-type or ΔCD24 cells with M2-like macrophages expressing Siglec-10 revealed that *CD24* genetic deletion alone was sufficient to potentiate phagocytosis (Fig. 2b). ΔCD24 cells were also significantly more sensitive to CD47 blockade (using Clone 5F9-G4[19]) than were wild-type cells, suggesting the cooperativity of combinatorial blockade of CD24 and CD47. To measure phagocytic clearance by automated live-cell microscopy, GFP^+ wild-type and ΔCD24 cells were labelled with the pH-sensitive dye pHrodo Red[20] and were co-cultured with macrophages. Over the course of 36 h, we found that ΔCD24 cells were more readily engulfed and degraded in the low-pH phagolysosome, as compared with wild-type cells (Fig. 2c).

The blockade of Siglec-10 using monoclonal antibodies augmented the phagocytic ability of macrophages, thereby confirming a role for Siglec-10 in inhibiting phagocytosis (Fig. 2d). To further investigate the effect of Siglec-10 expression on phagocytosis, we knocked out the *SIGLEC10* gene in donor-derived macrophages. Three days after electroporation with a single-guide RNA targeting the *SIGLEC10* locus, we observed a marked reduction in Siglec-10 expression relative to cells electroporated with Cas9 alone (Cas9 control) (Fig. 2e). *SIGLEC10* knockout macrophages demonstrated significantly greater phagocytic ability than Cas9 control macrophages (Fig. 2f).

Siglec-10 has been reported to interact with the highly sialylated form of CD24[13,14]. Accordingly, we observed that binding of Siglec-10–Fc (Fc, crystallizable fragment) to MCF-7 cells was considerably reduced upon surface desialylation (Fig. 2g, Extended Data Fig. 3b). This suggests that Siglec-10 has the capacity to recognize both protein and sialic acid ligands, and therefore probably has varied ligands that extend beyond CD24. Indeed, we observed that *CD24* deletion alone is insufficient to completely abrogate Siglec-10–Fc binding in the presence of surface sialylation (Extended Data Fig. 3a, b). However, in the absence of surface sialylation, Siglec-10–Fc binding was nearly abolished by *CD24* deletion, suggesting that CD24 is the primary protein ligand for Siglec-10 (Fig. 2h, Extended Data Fig. 3b). We found that desialylation did not reduce the enhancement of phagocytosis that was observed upon *CD24* deletion, indicating that CD24 sialylation is not required to inhibit phagocytosis (Extended Data Fig. 3c). Neither recombinant Siglec-5–Fc nor Siglec-9–Fc were found to bind $CD24^+$ MCF-7 cells, although both were highly expressed by donor-derived macrophages (Extended Data Fig. 3d–g).

To investigate the human therapeutic potential of these findings, we examined whether direct monoclonal antibody (mAb) blockade of CD24 could enhance the phagocytosis of $CD24^+$ human cancers by disrupting CD24–Siglec-10 signalling (Extended Data Fig. 4a). Automated live-cell microscopy revealed that MCF-7 pHRodo Red^+ cells treated with a CD24-blocking mAb (clone SN3)[21] were more readily engulfed into the low pH phagolysosome, as demonstrated by an enhanced red signal over time (Fig. 2i, Extended Data Fig. 4b). Substantial whole-cell phagocytosis was observed by confocal microscopy upon treatment with anti-CD24 mAb, and dual blockade of both CD24 and CD47 further augmented cellular engulfment (Extended Data Fig. 4c, d). Similarly, FACS-based measurements revealed a robust increase in phagocytosis upon the addition of anti-CD24 mAb as compared to the IgG control, which was greater than the effect observed with CD47 blockade (Fig. 3a; the gating strategy for in vitro phagocytosis is shown in Extended Data Fig. 5a). The response to anti-CD24 mAb was found to be dose-dependent and saturable (Extended Data Fig. 5b). CD24 blockade augmented the phagocytosis of all CD24-expressing cancer cell lines tested—including breast cancer (MCF-7), pancreatic

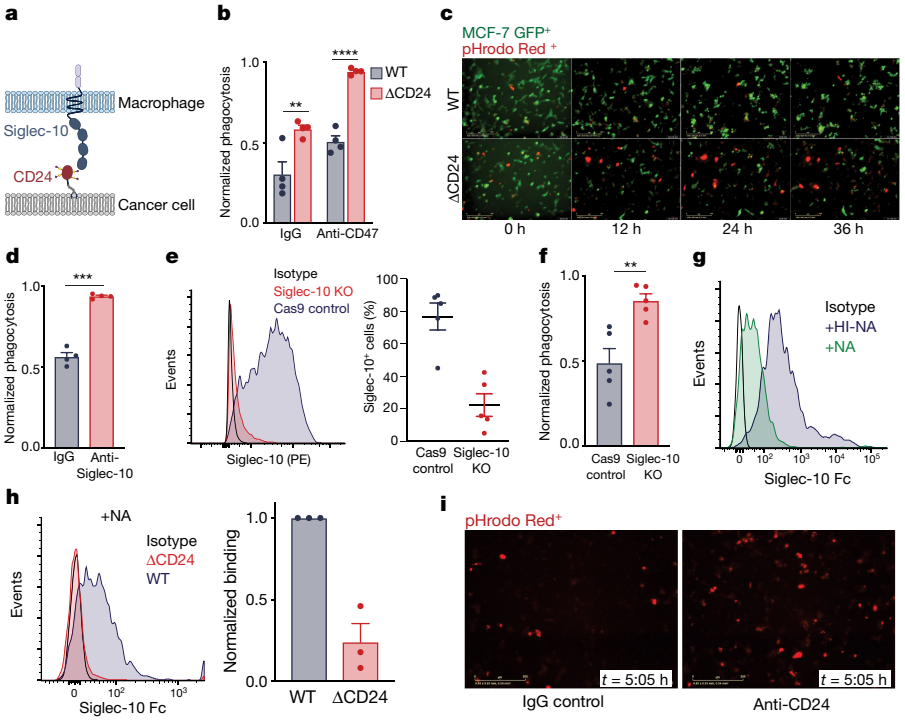


Fig. 2 | CD24 directly protects cancer cells from phagocytosis by macrophages. a, Schematic depicting interactions between macrophage-expressed Siglec-10 and CD24 expressed by cancer cells. **b**, Phagocytosis of $CD24^+$ MCF-7 cells (wild-type, WT) and $CD24^-$ (ΔCD24) MCF-7 cells, in the presence or absence of anti-CD47 mAb ($n = 4$ donors; two-way ANOVA with multiple comparisons correction, cell line $F_{(1,12)} = 65.65$; treatment $F_{(1,12)} = 40.30$, $^{**}P = 0.0045$, $^{****}P < 0.0001$). **c**, Representative phagocytosis images of pHrodo Red^+GFP^+ MCF-7 cells (wild-type, top; ΔCD24, bottom) over time; images are representative of two donors. **d**, Phagocytosis of wild-type MCF-7 cells, in the presence of anti-Siglec-10 mAb or IgG control ($n = 4$ donors; paired, two-tailed Student's *t*-test, $^{***}P = 0.0010$). **e**, Left, FACS-based measurement of Siglec-10 expression (phycoerythrin (PE)-conjugated) by Siglec-10 knockout (KO) macrophages (red) compared with Cas9 control (blue); right, frequency of $Siglec\text{-}10^+$ macrophages among Cas9 control compared with Siglec-10 knockout macrophages. Data are mean ± s.e.m. of $n = 5$ donors. **f**, Phagocytosis of wild-type MCF-7 cells by either Siglec-10 knockout or Cas9 control macrophages. Data are mean ± s.e.m. of $n = 5$ donors; paired, one-tailed Student's *t*-test, $^{**}P = 0.0035$. **g**, Flow-cytometry-based measurement of the binding of recombinant Siglec-10–Fc to MCF-7 wild-type cells treated with neuraminidase (+NA) or heat-inactivated neuraminidase (+HI-NA); plot is representative of two experimental replicates. **h**, Left, flow-cytometry-based measurement of the binding of Siglec-10–Fc to neuraminidase-treated MCF-7 wild-type cells compared with neuraminidase-treated MCF-7(ΔCD24) cells. Plot is representative of three experimental replicates. Right, normalized binding of Siglec-10–Fc to neuraminidase-treated MCF-7(ΔCD24) cells compared with neuraminidase-treated MCF-7 wild-type cells. Data are representative of three experimental replicates. **i**, Representative images from live-cell microscopy phagocytosis assays of pHrodo Red^+ MCF-7 cells treated with anti-CD24 mAb (right) or IgG control (left) at a time, *t*, of 5:05 h; images are representative of two donors and two experimental replicates.

adenocarcinoma (Panc1), pancreatic neuroendocrine tumour (APL1) and small-cell lung cancer (NCI-H82)—and no effect was observed with $CD24^-$ cells (U-87 MG) (Fig. 3b, Extended Data Fig. 5c). Upon dual treatment with CD24- and CD47-blocking antibodies, the induction of phagocytosis was increased to levels nearly 30 times that of the baseline in some cancers. Although genetic deletion of *CD47* alone did not alter the phagocytic susceptibility of MCF-7 cells, upon treatment with anti-CD24 mAb, ΔCD47 cells were more readily engulfed than were wild-type cells (Extended Data Fig. 5d). Dual treatment of pancreatic adenocarcinoma cells with anti-CD24 mAb and cetuximab enhanced phagocytosis relative to either treatment alone, demonstrating a potential synergy between anti-CD24 mAb and anti-solid-tumour mAbs (Extended Data Fig. 5e). An isotype-matched antibody against epithelial cellular adhesion molecule (EpCAM)—a surface marker that is highly expressed by MCF-7 cells—led to a modest increase in phagocytosis as compared to treatment with anti-CD24 mAb, which indicates that the vast majority of the observed increase in phagocytosis upon the addition of anti-CD24 mAb is due to loss of CD24 signalling and not due to Fc-mediated opsonization (Extended Data Fig. 6a). Both M2-like and M0 macrophages were found to respond equally to opsonization by anti-EpCAM antibodies (Extended Data Fig. 6b). Disruption of the interaction between the Fc portion of the anti-CD24 mAb and the Fc receptors—CD16 and CD32—led to a modest reduction in anti-CD24 mAb-induced phagocytosis, confirming that the Fc-mediated pro-phagocytic effect of the anti-CD24 mAb is minor (Extended Data Fig. 6c).

All Siglec-10-expressing macrophages responded to CD24 blockade (Extended Data Fig. 6d), and the magnitude of this response trended towards a correlation with Siglec-10 expression (Extended Data Fig. 6e). Genetic deletion of *SIGLEC10* led to a marked reduction in the response to CD24 blockade, which indicates that anti-CD24 mAb specifically disrupts CD24–Siglec-10 signalling (Fig. 3c). Expression of CD24 correlated with response to CD24 blockade as well as with baseline phagoytosis levels, suggesting that tissue-specific expression of CD24 is a dominant 'don't eat me' signal and highlighting the potential value of CD24 expression as a predictor of the innate anti-tumour immune response (Fig. 3d, Extended Data Fig. 6f).

Ovarian cancer cells were collected from patients with metastatic ovarian cancer and were treated with anti-CD24 mAb in order to measure phagocytosis of primary human tumours. (Fig. 3e). In these cases, CD24 blockade yielded a significantly greater effect than CD47 blockade, and dual treatment with both CD24- and CD47-blocking antibodies augmented phagocytosis at least additively (Fig. 3f). Furthermore, treatment of primary human TNBC cells with anti-CD24 mAb promoted phagocytic clearance by macrophages, whereas in these cases CD47 blockade had no effect on phagocytosis; this indicates that anti-CD24 mAb may be efficacious in cancers that show resistance to CD47 blockade (Extended Data Fig. 6g).

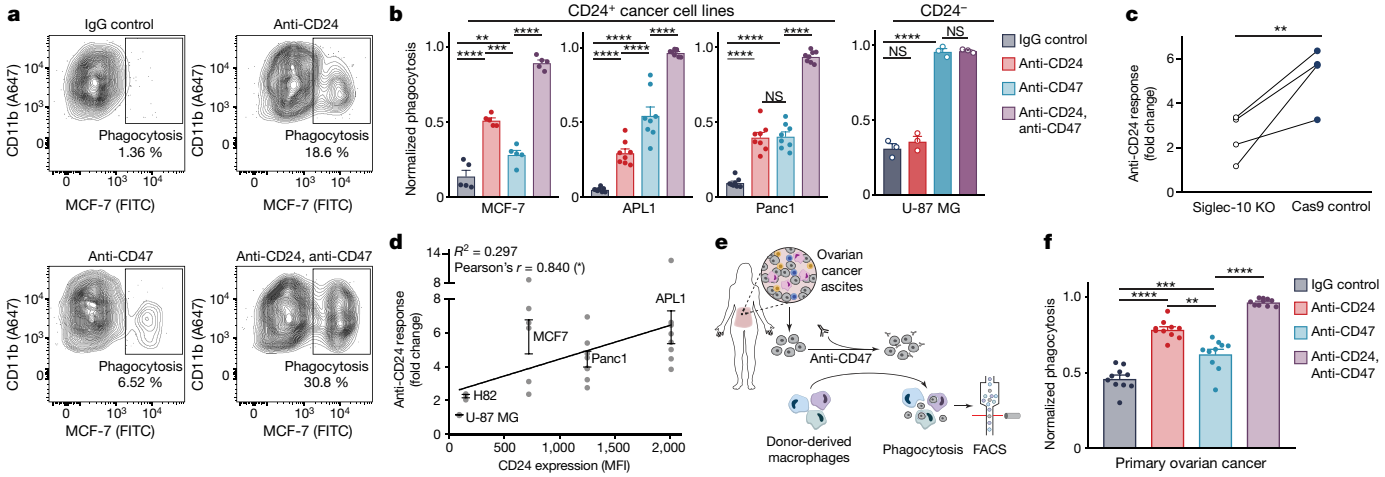


Fig. 3 | Treatment with anti-CD24 mAb promotes phagocytic clearance of human cancer cells. a, Representative flow-cytometry plots depicting the phagocytosis of MCF-7 cells treated with anti-CD24 mAb, CD47 mAb or dual treatment, compared with the IgG control. Plots are representative of five donors. FITC, fluorescein isothiocyanate. **b**, Phagocytosis of MCF-7 ($n = 5$ donors), APL1 ($n = 8$ donors) and Panc1 ($n = 8$ donors) cell lines (left) and of the U-87 GM cell line ($n = 3$ donors; solid bars) (right) in the presence of anti-CD24 mAb, anti-CD47 mAb or dual treatment, compared with IgG control (one-way ANOVA with multiple comparisons correction; MCF-7 $F_{(3,16)} = 145.6$, APL1 $F_{(3,28)} = 144.7$, Panc1 $F_{(3,28)} = 220.7$, U-87 MG $F_{(3,8)} = 200.4$; NS, not significant; $^{**}P = 0.0092$, $^{***}P = 0.0001$, $^{****}P < 0.0001$). **c**, Response to anti-CD24 mAb treatment by Siglec-10 knockout compared with Cas9 control macrophages ($n = 4$ donors, connecting lines indicate matched donor. Paired, one-tailed Student's t-test, $^{**}P = 0.0089$). **d**, Pearson correlation between CD24 expression (x axis) and mean anti-CD24 mAb response (y axis) (n values are the same as those listed in **b** and Extended Data Fig. 5c. Linear regression is shown. Data are mean ± s.e.m. $^{*}P = 0.0375$). MFI, median fluorescence intensity. **e**, Workflow to measure the phagocytosis of primary ovarian cancer. **f**, Phagocytosis of primary ovarian cancer cells treated with anti-CD24 mAb, anti-CD47 mAb or dual treatment, compared with IgG control ($n = 10$ macrophage donors, $n = 1$ primary ovarian cancer ascites donor) (one-way ANOVA with multiple comparisons correction, $F_{(2.110,\ 18.99)} = 121.5$, $^{**}P = 0.0078$, $^{***}P = 0.0006$, $^{****}P < 0.0001$). Data are mean ± s.e.m.

To investigate whether the protection against phagocytosis conferred by CD24 could be recapitulated in vivo, GFP-luciferase^{+} MCF-7 wild-type or MCF-7(ΔCD24) cells were engrafted into NOD.Cg-Prkdc^{SCID}Il2rg^{tm1Wjl}/SzJ (NSG) mice[22]. Three weeks after engraftment, we found that CD24-deficient tumours exhibited augmented levels of in vivo phagocytosis by infiltrating TAMs as compared to wild-type tumours, and TAMs that infiltrated the CD24-deficient tumours were also of a more inflammatory phenotype (Extended Data Figs. 7, 8a, b). Over weeks, we observed a robust reduction in the growth of ΔCD24 tumours as compared to wild-type tumours (Fig. 4a, b). Notably, the sublines assessed had no measurable cell-autonomous differences in proliferation in vitro (Extended Data Fig. 8c). After 35 days of growth, the polyclonal ΔCD24 tumours had become largely CD24^{+}, which is consistent with the selection against CD24^{-} cells by TAMs and the

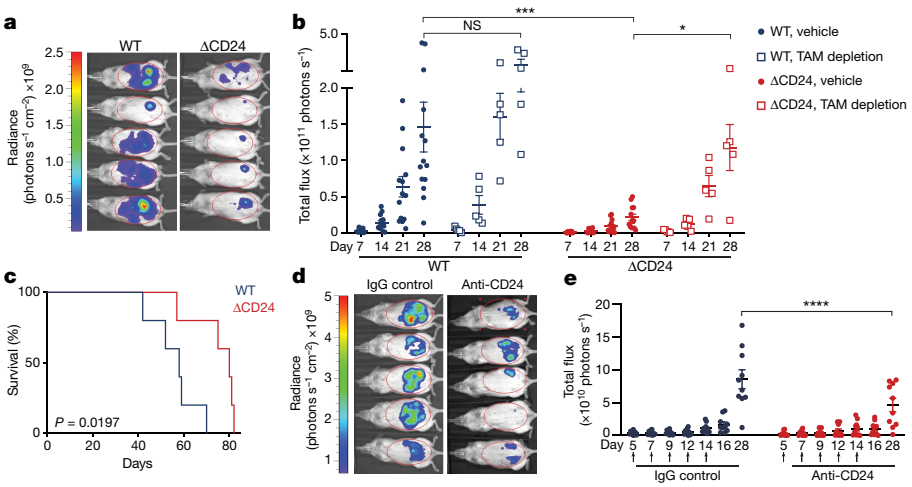


Fig. 4 | CD24 protects cancer cells from macrophage attack in vivo. a, Representative bioluminescence images of day-21 tumours in mice engrafted with MCF-7 wild-type and MCF-7(ΔCD24) tumours. Image representative of two independent experimental cohorts. **b**, Burden of MCF-7 wild-type compared with MCF-7(ΔCD24) tumours in mice with TAMs (vehicle) or in TAM-depleted mice (treated with anti-CSF1R) as measured by bioluminescence (WT vehicle, $n = 14$; WT TAM depletion, $n = 5$; ΔCD24 vehicle, $n = 13$; ΔCD24 TAM depletion, $n = 5$. Two-way ANOVA with multiple comparisons correction, tumour genotype $F_{(3,33)} = 11.75$, $^{*}P = 0.0296$, $^{***}P = 0.0009$). **c**, Survival analysis of vehicle-treated mice in **b**; P value computed by a log-rank (Mantel–Cox) test (WT, $n = 5$; ΔCD24, $n = 5$). **d**, Representative bioluminescence image of day-33 tumours in mice with MCF-7 tumours treated with either IgG control or anti-CD24 mAb (image representative of two experimental cohorts). Data are mean ± s.e.m. **e**, Burden of MCF-7 wild-type tumours treated with IgG control (blue) and anti-CD24 mAb (red) as measured by bioluminescence (IgG, $n = 10$; anti-CD24 mAb, $n = 10$. The days on which the treatments were administered are indicated by arrows. Data from two experimental cohorts. Two-way ANOVA with multiple-comparisons correction, tumour treatment $F_{(1,126)} = 5.679$, $^{****}P < 0.0001$).

emergence of subclones of $CD24^+$ cells that did not have biallelic *CD24* deletion (Extended Data Fig. 8d, e). TAM depletion did not significantly alter the burden of wild-type tumours, whereas the loss of TAMs largely abrogated the reduction of tumour growth that was observed in ΔCD24 tumours, indicating that increased TAM-mediated clearance of ΔCD24 cells was responsible for the diminished tumour burden (Fig. 4b, Extended Data Fig. 8f). This reduction in tumour growth—attributed to enhanced phagocytic clearance—resulted in a significant survival advantage for mice engrafted with CD24-deficient tumours (Fig. 4c).

To determine whether the mouse homologue of human CD24—gene name *Cd24a*—could similarly confer protection against phagocytic clearance of cancer cells, we generated a subline of the mouse epithelial ovarian cancer line ID8 that lacked CD24 (ID8ΔCd24a). Wild-type or ΔCd24a cells expressing GFP were injected intraperitoneally into NSG mice. After one week of growth, we observed that loss of *Cd24a* was sufficient to significantly promote in vivo phagocytosis by NSG macrophages (Extended Data Fig. 9a). To assess the effect of mouse CD24 in a syngeneic, fully immunocompetent setting, ID8 wild-type or ID8ΔCd24a cells were engrafted intraperitoneally into C57Bl/6J mice. We observed that loss of CD24 was sufficient to substantially reduce tumour growth over several weeks (Extended Data Fig. 9b, c).

To demonstrate that the enhancement of anti-tumour immunity could be modulated by therapeutic blockade of CD24, NSG mice with established MCF-7 wild-type tumours were treated with anti-CD24 monoclonal antibody for two weeks. Anti-CD24 therapy resulted in significant reduction of tumour growth compared to the IgG-treated control (Fig. 4d, e, Extended Data Fig. 9d).

Potential off-target effects of anti-CD24 mAb treatment in humans include depletion of B cells, owing to high CD24 expression by B cells. Indeed, phagocytic clearance of healthy B cells was observed upon treatment with anti-CD24 mAb (Extended Data Fig. 10a). However, we found that—unlike anti-CD47 mAbs[4]—the anti-CD24 mAb demonstrated no detectable binding to human red blood cells, even though mouse CD24a is expressed by mouse red blood cells (Extended Data Fig. 10b).

CD24 is a potent anti-phagocytic, 'don't eat me' signal that is capable of directly protecting cancer cells from attack by Siglec-10-expressing macrophages. Monoclonal antibody blockade of CD24–Siglec-10 signalling robustly enhances clearance of $CD24^+$ tumours, which indicates promise for CD24 blockade in immunotherapy. Both ovarian[23] and breast cancer have demonstrated weaker responses to anti-PD-L1/PD-1 immunotherapies than have other cancers[24–26], which suggests that an alternative strategy may be required to achieve responses across a wide range of cancers. It is notable that the 'don't eat me' signals CD47, PD-L1, B2M—and now CD24—each involve macrophage signalling based on immunoreceptor-tyrosine-based inhibition motifs. This may indicate a conserved mechanism that leads to immunoselection of the subset of macrophage-resistant cancer cells, resulting in tumours that—by nature—avoid macrophage surveillance and clearance. CD24 expression may provide immediate predictive value of the responsiveness of tumours to existing immunotherapies, in that high CD24 expression may inhibit response to therapies that are reliant on macrophage function. Expression of CD24 and CD47 was found to be inversely related among patients with diffuse large B cell lymphoma (Extended Data Fig. 10c). The percentage of patients with CD24 overexpression compares well with the response rates observed with anti-CD47 + rituximab combination therapy in this disease[4], opening up the possibility that particular tumours might respond differentially to treatment with anti-CD24 and/or anti-CD47 mAbs. Determining the collective expression of pro- and anti-phagocytic signals expressed by cancers and associated macrophages could enable better prediction of which patients may respond to treatment. This work defines CD24–Siglec-10 as an innate immune checkpoint that is essential for mediating anti-tumour immunity, and provides evidence for the therapeutic potential of CD24 blockade, with particular promise for the treatment of ovarian and breast cancers.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1456-0.

Received: 24 January 2019; Accepted: 2 July 2019;
Published online 31 July 2019.

1. Majeti, R. et al. CD47 is an adverse prognostic factor and therapeutic antibody target on human acute myeloid leukemia stem cells. *Cell* **138**, 286–299 (2009).
2. Gordon, S. R. et al. PD-1 expression by tumour-associated macrophages inhibits phagocytosis and tumour immunity. *Nature* **545**, 495–499 (2017).
3. Barkal, A. A. et al. Engagement of MHC class I by the inhibitory receptor LILRB1 suppresses macrophages and is a target of cancer immunotherapy. *Nat. Immunol.* **19**, 76–84 (2018).
4. Advani, R. et al. CD47 blockade by Hu5F9-G4 and rituximab in non-Hodgkin's lymphoma. *N. Engl. J. Med.* **379**, 1711–1721 (2018).
5. Willingham, S. B. et al. The CD47-signal regulatory protein alpha (SIRPa) interaction is a therapeutic target for human solid tumors. *Proc. Natl Acad. Sci. USA* **109**, 6662–6667 (2012).
6. Pirruccello, S. J. & LeBien, T. W. The human B cell-associated antigen CD24 is a single chain sialoglycoprotein. *J. Immunol.* **136**, 3779–3784 (1986).
7. Chen, G. Y., Brown, N. K., Zheng, P. & Liu, Y. Siglec-G/10 in self–nonself discrimination of innate and adaptive immunity. *Glycobiology* **24**, 800–806 (2014).
8. Chen, W. et al. Induction of Siglec-G by RNA viruses inhibits the innate immune response by promoting RIG-I degradation. *Cell* **152**, 467–478 (2013).
9. Chen, G. Y. et al. Amelioration of sepsis by inhibiting sialidase-mediated disruption of the CD24–SiglecG interaction. *Nat. Biotechnol.* **29**, 428–435 (2011).
10. Chen, G. Y., Tang, J., Zheng, P. & Liu, Y. CD24 and Siglec-10 selectively repress tissue damage-induced immune responses. *Science* **323**, 1722–1725 (2009).
11. Toubai, T. et al. Siglec-G–CD24 axis controls the severity of graft-versus-host disease in mice. *Blood* **123**, 3512–3523 (2014).
12. Crocker, P. R., Paulson, J. C. & Varki, A. Siglecs and their roles in the immune system. *Nat. Rev. Immunol.* **7**, 255–266 (2007).
13. Abram, C. L. & Lowell, C. A. Shp1 function in myeloid cells. *J. Leukoc. Biol.* **102**, 657–675 (2017).
14. Dietrich, J., Cella, M. & Colonna, M. Ig-like transcript 2 (ILT2)/leukocyte Ig-like receptor 1 (LIR1) inhibits TCR signaling and actin cytoskeleton reorganization. *J. Immunol.* **166**, 2514–2521 (2001).
15. Tarhriz, V. et al. Overview of CD24 as a new molecular marker in ovarian cancer. *J. Cell. Physiol.* **234**, 2134–2142 (2019).
16. Kristiansen, G. et al. CD24 expression is a new prognostic marker in breast cancer. *Clin. Cancer Res.* **9**, 4906–4913 (2003).
17. Karaayvaz, M. et al. Unravelling subclonal heterogeneity and aggressive disease states in TNBC through single-cell RNA-seq. *Nat. Commun.* **9**, 3588 (2018).
18. Mantovani, A., Sozzani, S., Locati, M., Allavena, P. & Sica, A. Macrophage polarization: tumor-associated macrophages as a paradigm for polarized M2 mononuclear phagocytes. *Trends Immunol.* **23**, 549–555 (2002).
19. Liu, J. et al. Pre-clinical development of a humanized anti-CD47 antibody with anti-cancer therapeutic potential. *PLoS ONE* **10**, e0137345 (2015).
20. Miksa, M., Komura, H., Wu, R., Shah, K. G. & Wang, P. A novel method to determine the engulfment of apoptotic cells by macrophages using pHrodo succinimidyl ester. *J. Immunol. Methods* **342**, 71–77 (2009).
21. Okazaki, M., Luo, Y., Han, T., Yoshida, M. & Seon, B. K. Three new monoclonal antibodies that define a unique antigen associated with prolymphocytic leukemia/non-Hodgkin's lymphoma and are effectively internalized after binding to the cell surface antigen. *Blood* **81**, 84–94 (1993).
22. Shultz, L. D. et al. Human lymphoid and myeloid cell development in NOD/LtSz-scid IL2R γ null mice engrafted with mobilized human hemopoietic stem cells. *J. Immunol.* **174**, 6477–6489 (2005).
23. Varga A. et al. Pembrolizumab in patients (pts) with PD-L1–positive ($PD\text{-}L1^+$) advanced ovarian cancer: updated analysis of KEYNOTE-028. *J. Clin. Oncol.* **35**, 5513 (2017).
24. Nanda, R. et al. Pebrolizumab in patients with advanced triple-negative breast cancer: Phase Ib KEYNOTE-012 study. *J. Clin. Oncol.* **34**, 2460–2467 (2016).
25. Alsaab, H. O. et al. PD-1 and PD-L1 checkpoint signaling inhibition for cancer immunotherapy: mechanism, combinations, and clinical outcome. *Front. Pharmacol.* **8**, 561 (2017).
26. Ayers M. *et al.* Molecular profiling of cohorts of tumor samples to guide clinical development of pembrolizumab as monotherapy. *Clinical Cancer Res.* https://doi.org/10.1158/1078-0432.CCR-18-1316 (2018).

METHODS

Statistics. Sample sizes were modelled after those from existing publications regarding in vitro immune killing assays and in vivo tumour growth assays, and an independent statistical method was not used to determine sample size. Statistical tests were performed in GraphPad Prism 8.

Human tumour bulk RNA-sequencing analysis. RNA-sequencing data regarding expression levels for *CD24*, *CD274* (*PD-L1*), *CD47* and *B2M* from human tumours and matched healthy tissues collected by TCGA, TARGET, and the Genotype-Tissue Expression Project (GTEX) were downloaded as $\log_2$(normalized counts +1) values from UCSC Xena[27] (https://xenabrowser.net/) with the query 'TCGA TARGET GTEX'. Tumour types were filtered for those with ≥9 individual patients for either tumour or healthy tissues. For instances in which there existed both TCGA-matched healthy tissues and GTEX healthy tissues, all healthy tissues were combined for analyses. Abbreviations for TCGA studies and number of samples analysed are listed in Supplementary Table 1. Survival analysis was performed by stratifying patients into high or low *CD24* expression using median expression values, and Kaplan–Meier plots were generated and analysed using Prism 8. Two-dimensional contour plots were generated using Plotly (Plotly Technologies)

Single-cell RNA-sequencing analysis. Raw files from previously sequenced TNBC (accession number PRJNA485423) were downloaded from the NCBI Sequence Read Archive (ref. 17). The 1,539 single-cell RNA-sequencing data was aligned to the human genome (GRCh38) using STAR (version 2.5.3a) and gene counts (gene models from ENSEMBL release 82) were determined using htseq-count (intersection-nonempty mode, secondary and supplementary alignments ignored, no quality score requirement). The expression matrix was transformed to gene counts per million (c.p.m.) sequenced reads for each cell. High-quality cells were defined as those that had at least 200,000 c.p.m. and at least 500 genes expressed. This resulted in 1,001 cells.

Marker genes used in ref. 17 were used to determine cell types. This was done using UMAP (nonlinear dimensionality reduction algorithm) on log-transformed c.p.m. values for the marker genes and labelling each of the five clusters identified on the basis of which cell markers were most expressed (see Extended Data Fig. 1d). Scatter plots were constructed using this UMAP transformation with colouring as described in the figure legends.

Cell culture. All cell lines were purchased from the American Type Culture Collection (ATCC) with the exception of the APL1 cells, which were a gift from G. Krampitz (MD Anderson), and the ID8 cells, which were obtained from the laboratory of O.D. The human NCI-H82 and APL1 cells were cultured in RPMI+GlutaMax (Life Technologies) + 10% fetal bovine serum (FBS) + 100 U ml^{-1} penicillin/streptomycin (Life Technologies). Cell lines were not independently authenticated beyond the identity provided from the ATCC. The human MCF-7, Panc1 and U-87 GM cell lines were cultured in DMEM+GlutaMax + 10% FBS + 100 U ml^{-1} penicillin/streptomycin. The murine ovarian carcinoma cell line, ID8, was cultured in DMEM + 4% FBS + 10% insulin/transferrin/selenium (Corning) + 100 U ml^{-1} penicillin/streptomycin. All cells were cultured in a humidified, 5% CO_2 incubator at 37 °C. All cell lines were tested for mycoplasma contamination.

Generation of MCF-7 and ID8 sub-lines. Parental MCF-7 and ID8 were infected with GFP–luciferase lentivirus in order to generate MCF-7-GFP-luc$^+$ and ID8-GFP-luc$^+$ cell lines, respectively. After 48 h, cells were collected and sorted by FACS in order to generate pure populations of GFP$^+$ cells. The MCF-7ΔCD24-GFP-luc$^+$ and ID8ΔCd24a-GFP-luc$^+$ sub-lines were generated by electroporating cells with recombinant CRISPR–Cas9 ribonucleoprotein (RNP), as described previously[5]. In brief, CRISPR–Cas9 guide RNA molecules targeting human *CD24* and mouse *Cd24a*, respectively, were purchased as modified, hybridized RNA molecules (Synthego) and assembled with Cas9-3NLS nuclease (IDT) via incubation at 37 °C for 45 min. Next, 2×10^6 MCF-7-GFP-luc$^+$ or ID8-GFP-luc$^+$ cells were collected, combined with corresponding complexed Cas9/RNP and electroporated using the Lonza Nucleofector IIb using Kit V (VCA-1003). After 48 h of culture, genetically modified cells were collected and purified through at least three successive rounds of FACS sorting in order to generate pure cell lines. Sequences for the single-guide RNA (sgRNA) molecules used are as follows: human *CD24* sgRNA: CGGUGCGCGGCGCGUCUAGC; *hCD47* sgRNA: AAUAGUAGCUGAGCUGAUCC; and mouse *Cd24a* sgRNA: AUAUUCUGGUUACCGGGAAA.

In vitro cell proliferation assay. Proliferation of the MCF-7 wild-type and MCF-7ΔCD24 cell lines was measured with live-cell microscopy using an Incucyte (Sartorius). Cells were each plated at around 10% confluence. Percentage confluence after cell growth was measured as per the manufacturer's instructions every 8 h for 64 h.

Neuraminidase treatment and recombinant Siglec-binding assay. MCF-7 cells were treated with either neuraminidase (from *Vibrio cholerae*, Roche) (1×10^6 cells per 100 U per ml or neuraminidase that was heat-inactivated for 15 min at 95 °C before incubation for 1 h at 37 °C in serum-free medium, after which reactions were quenched with serum before analysis. Recombinant Siglecs (10, 5 and 9) were purchased as human Fc-fusion proteins from R&D Systems. Binding of recombinant Siglecs versus human IgG1 control was assayed at a concentration of 1×10^5 cells per mg per ml at 37 °C for 1 h, in the absence of EDTA. Cells were stained with a fluorescently conjugated anti-human Fc antibody (BioLegend) to enable the measurement of recombinant Siglec binding by flow cytometry.

Macrophage generation and stimulation. Primary human donor-derived macrophages were generated as described previously[28]. In brief, leukocyte reduction system chambers from anonymous donors were obtained from the Stanford Blood Center. Peripheral monocytes were purified through successive density gradients using Ficoll (Sigma-Aldrich) and Percoll (GE Healthcare). Monocytes were then differentiated into macrophages by 7–9 days of culture in IMDM + 10% AB human serum (Life Technologies). Unless otherwise stated, macrophages used for all in vitro phagocytosis assays were stimulated with 50 ng ml^{-1} human TGFβ1 (Roche) and 50 ng ml^{-1} human IL-10 (Roche) on days 3–4 of differentiation until use on days 7–9. IL-4 stimulation was added at a concentration of 20 ng ml^{-1} on days 3–4 of differentiation until use on days 7–9.

Human macrophage knockouts. Genetic knockouts in primary human donor-derived macrophages were performed as described previously[5]. In brief, sgRNA molecules targeting the first exon of *SIGLEC10* were purchased from Synthego as modified, hybridized RNA molecules. The *SIGLEC10* sgRNA sequence used is: AGAAUCUCCCAUCCAUAGCC. Mature (day 7) donor-derived macrophages were electroporated with Cas9 ribonuclear proteins using the P3 Primary Cell Nucleofection Kit (Lonza V4XP-3024). Macrophages were collected for analysis and functional studies 72 h after electroporation. Indel frequencies were quantified using TIDE software as described previously[29].

Human samples. The Human Immune Monitoring Center Biobank, the Stanford Tissue Bank, O.D. and G. Wernig all received IRB approval from the Stanford University Administrative Panels on Human Subjects Research and complied with all ethical guidelines for human subjects research to obtain samples from patients with ovarian cancer and breast cancer, and received informed consent from all patients. Single-cell suspensions of solid tumour specimens were attained by mechanical dissociation using a straight razor, followed by an enzymatic dissociation in 10 ml of RPMI + 10 μg ml^{-1} DNaseI (Sigma-Aldrich) + 25 μg ml^{-1} Liberase (Roche) for 30–60 min at 37 °C with vigorous pipetting every 10 min to promote dissociation. After a maximum of 60 min, dissociation reactions were quenched with 4 °C RPMI + 10% FBS, filtered through a 100-μm filter and centrifuged at 400*g* for 10 min at 4 °C. Red blood cells in samples were then lysed by resuspending the tumour pellet in 5 ml ACK Lysing Buffer (Thermo Fisher Scientific) for 5 min at room temperature. Lysis reactions were quenched by the addition of 20 ml RPMI + 10% FBS, and samples were centrifuged at 400*g* for 10 min at 4 °C. Samples were either directly analysed, or resuspended in Bambanker (Wako Chemicals), aliquoted into cryovials and frozen before analysis.

FACS of primary human tumour samples. Single-cell suspensions of primary human tumour samples were obtained (described above), and frozen samples were thawed for 3–5 min at 37 °C, washed with DMEM + 10% FBS, and centrifuged at 400*g* for 5 min at 4 °C. Samples were then resuspended in FACS buffer at a concentration of 1 million cells per ml and blocked with monoclonal antibody to CD16/32 (Trustain fcX, BioLegend) for 10–15 min on ice before staining with antibody panels. Antibody panels are listed, with clones, fluorophores, usage purpose, and concentrations used in Supplementary Table 2. Samples were stained for 30 min on ice, and subsequently washed twice with FACS buffer and resuspended in buffer containing 1 μg ml^{-1} DAPI before analysis. Fluorescence compensations were performed using single-stained UltraComp eBeads (Affymetrix). Gating for immune markers and DAPI was performed using fluorescence minus one controls, while CD24$^+$ and Siglec-10$^+$ gates were drawn on the basis of appropriate isotype controls (see Extended Data Fig. 2a for gating strategy). Flow cytometry was performed either on a FACSAria II cell sorter (BD Biosciences) or on an LRSFortessa Analyzer (BD Biosciences) and all flow cytometry data reported in this work was analysed using FlowJo. Human tumour gating schemes were as follows: human TAMs: DAPI$^-$, EpCAM$^-$, CD14$^+$, CD11b$^+$; human tumour cells: DAPI$^-$, CD14$^-$, EpCAM$^+$.

Flow-cytometry-based phagocytosis assay. All in vitro phagocytosis assays reported here were performed by co-culture target cells and donor-derived macrophages at a ratio of 100,000 target cells to 50:000 macrophages for 1–2 h in a humidified, 5% CO_2 incubator at 37 °C in ultra-low-attachment 96-well U-bottom plates (Corning) in serum-free IMDM (Life Technologies). Cells with endogenous fluorescence were collected from plates using TrypLE Express (Life Technologies) before co-culture. Cells from cell lines that lack endogenous fluorescence—NCI-H82 and Panc1—were collected using TrypLE Express and fluorescently labelled with Calcein AM (Invitrogen) by suspending cells in PBS + 1:30,000 Calcein AM as per the manufacturer's instructions for 15 min at 37 °C and washed twice with 40 ml PBS before co-culture. For TNBC primary-sample phagocytosis assays, tumours were acquired fresh on the day of resection and dissociated as

described above. EpCAM$^+$ tumour cells were purified on an autoMACS pro separator (Miltenyi) by first depleting samples of myeloid cells using anti-CD14 microbeads (Miltenyi, 1:50) followed by an enrichment with anti-EpCAM microbeads (Miltenyi, 1:50). For primary ovarian cancer ascites assays, ovarian ascites samples were frozen as described above, thawed and directly labelled with Calcein-AM (Invitrogen) at a concentration of 1:30,000. For primary B cell phagocytosis assays, B cells were enriched from pooled donor peripheral blood mononuclear cell (PBMC) fractions using an autoMACS pro separator (Miltenyi) using anti-CD19 microbeads (Miltenyi, 1:50). For Fc-receptor blockade phagocytosis assays, macrophages were pre-treated with 10 μg ml^{-1} human Fc-receptor blocking solution (BioLegend) for 45 min at 4 °C, and subsequent co-culture with mAb-treated target cells was conducted in the presence of 10 μg ml^{-1} human Fc-receptor blocking solution. For all assays, macrophages were collected from plates using TrypLE Express. For phagocytosis assays involving treatment with monoclonal antibodies including anti-CD24 (Clone SN3, Novus Biologics) and anti-CD47 (Clone 5F9-G4, acquired from Forty Seven), all antibodies or appropriate isotype controls were added at a concentration of 10 μg ml^{-1}. After co-culture, phagocytosis assays were stopped by placing plates on ice, centrifuged at 400*g* for 5 min at 4 °C and stained with A647-labelled anti-CD11b (Clone M1/70, BioLegend) to identify human macrophages. Assays were analysed by flow cytometry on an LRSFortessa Analyzer (BD Biosciences) or a CytoFLEX (Beckman), both using a high-throughput auto-sampler. Phagocytosis was measured as the number of CD11b$^+$GFP$^+$ macrophages, quantified as a percentage of the total CD11b$^+$ macrophages. Each phagocytosis reaction (independent donor and experimental group) was performed in technical triplicate as a minimum, and outliers were removed using GraphPad Outlier Calculator (https://www.graphpad.com/quickcalcs/Grubbs1.cfm). To account for innate variability in raw phagocytosis levels among donor-derived macrophages, phagocytosis was normalized to the highest technical replicate per donor. All biological replicates indicate independent human macrophage donors. See Supplementary Table 2 for antibodies and isotype controls used in this study, and Extended Data Fig. 5a for example gating. Response to anti-CD24 mAb was computed by the fold change in phagocytosis between anti-CD24 mAb treatment and IgG control.

Time-lapse live-cell-microscopy-based phagocytosis assay. Non-fluorescently labelled MCF-7 cells were collected using TrypLE express and labelled with pHrodo Red succinimidyl ester (Thermo Fisher Scientific) as per the manufacturer's instructions at a concentration of 1:30,000 in PBS for 1 h at 37 °C, followed by two washes with DMEM + 10% FBS + 100 U ml^{-1} penicillin/streptomycin. Donor-derived macrophages were collected using TrypLE express and 50,000 macrophages were added to clear, 96-well flat-bottom plates and allowed to adhere for 1 h at 37 °C. After macrophage adherence, 100,000 pHrodo-Red-labelled MCF-7 cells + 10 μg ml^{-1} anti-CD24 antibody (SN3) were added in serum-free IMDM. The plate was centrifuged gently at 50*g* for 2 min in order to promote the timely settlement of MCF-7 cells into the same plane as adherent macrophages. Phagocytosis assay plates were then placed in an incubator at 37 °C and imaged at 10–20-min intervals using an Incucyte (Essen). The first image time point (reported as $t = 0$) was generally acquired within 30 min of co-culture. Images were acquired using a 20× objective at 800-ms exposures per field. Phagocytosis events were calculated as the number of pHrodo-red$^+$ events per well and values were normalized to the maximum number of events measured across technical replicates per donor. Thresholds for calling pHrodo-red$^+$ events were set on the basis of intensity measurements of pHrodo-red-labelled cells that lacked macrophages.

High-resolution phagocytosis microscopy. Fluorescently labelled MCF-7 cells (mCherry$^+$) and donor-derived macrophages were collected as described above. Suspensions consisting of 50,000 macrophages and 100,000 MCF-7 cells + 10 μg ml^{-1} antibody or isotype control in serum-free IMDM were placed into an untreated 24-well plate, in order to allow for adherence of donor-derived macrophages while preventing MCF-7 adherence. Reactions were incubated for 6 h in an incubator at 37 °C. After incubation, wells were washed vigorously five times with serum-free IMDM in order to wash away non-phagocytosed MCF-7 cells. Whole-cell phagocytosis was evaluated using a Leica DMI 6000B fluorescent microscope and an Olympus IX83. High-resolution *z*-stack images were taken on a Zeiss LSM800 confocal microscope. All images were processed in ImageJ and Adobe Illustrator.

Mice. NOD.Cg-*Prkdc*^{scid}*Il2rg*^{tm1Wjl}/SzJ (NSG) mice were obtained from in-house breeding stocks. C57Bl/6J mice were obtained from The Jackson Laboratory. All experiments were carried out in accordance with ethical care guidelines set by the Stanford University Administrative Panel on Laboratory Animal Care (APLAC). In compliance with Stanford APLAC protocol (26270), mice in long-term tumour studies were continually monitored to ensure adequate body condition scores and to ensure that tumours were less than 2.5 cm in diameter and that there was less than 50% ulceration. Female mice were used for all studies. Investigators were not blinded for animal studies.

In vivo phagocytosis analysis. For ID8 peritoneal phagocytosis analysis, 4 × 10^6, ID8-WT-GFP-luc$^+$ cells or ID8-ΔCd24a-GFP-luc$^+$ cells were engrafted into 6–8-week-old female NSG mice via intraperitoneal injection of single-cell suspensions in PBS. After 7 days, cells were collected by peritoneal lavage. For MCF-7 xenograft phagocytosis analysis, female NSG mice, 6–10 weeks of age, were engrafted with 4 × 10^6 MCF-7-WT-GFP-luc$^+$ cells or MCF-7- MCF-7-ΔCD24-GFP-luc$^+$ cells by injection of a single-cell suspension in 25% Matrigel Basement Membrane Matrix (Corning) + 75% RPMI orthotopically into the mammary fat pad. Tumours were allowed to grow for 28 days, after which tumours were resected and dissociated mechanically and enzymatically as described above. Single-cell suspensions of tumours were blocked using anti-CD16/32 (mouse TruStain FcX, BioLegend) for 15 min on ice as described above, before staining. Phagocytosis was measured as the percentage of CD11b$^+$F4/80$^+$ TAMs that were also GFP$^+$ (see Extended Data Fig. 7 for example gating). Mouse TAM gating schemes were as follows: mouse TAMs: DAPI$^-$, CD45$^+$, CD11b$^+$, F480$^+$; M1-like mouse TAMs: DAPI$^-$, CD45$^+$, CD11b$^+$, F480$^+$, CD80$^+$.

In vivo xenograft tumour-growth experiments. Female NSG mice, 6–10 weeks of age, were engrafted with 4 × 10^6 MCF-7-WT-GFP-luc$^+$ cells or MCF-7-ΔCD24-GFP-luc$^+$ cells as described above. Tumours were measured using bioluminescence imaging beginning 7 days post-engraftment and continuing every 7 days until day 28. Mice were injected intraperitoneally with firefly D-luciferin at 140 mg kg^{-1} in PBS and images were acquired 10 min after luciferin injection using an IVIS Spectrum (Perkin Elmer). Total flux was quantified using Living Image 4.0 software. For survival analyses, deaths were reported as the days on which the primary tumour burden reached 2.5 cm and/or the body condition scoring values fell below those allowed by our animal protocols.

In vivo macrophage depletion treatment study. Female NSG mice, 6–10 weeks of age, were depleted of macrophages as described previously[4] by treatment with 400 μg CSF1R antibody per mouse or PBS (vehicle) (BioXCell, Clone AFS98) three times per week for 18 days before engraftment, and throughout the duration of the experiment. Successful tissue resident macrophage depletion was confirmed by flow cytometry before tumour engraftment by peritoneal lavage and flow cytometry analysis (Extended Data Fig. 8f). Macrophage-depleted animals or vehicle treated animals were randomized before being engrafted with either MCF-7-WT-GFP-luc$^+$ or MCF-7-ΔCD24-GFP-luc$^+$ cells as described above.

Immunocompromised tumour treatment studies. Female NSG mice (6–8 weeks old) were engrafted with 4 × 10^6 MCF-7-WT-GFP-luc$^+$ cells. On day 5 after engraftment, the total flux of all tumours was measured using bioluminescence imaging and engraftment outliers were removed using GraphPad Outlier Calculator. Mice were randomized into treatment groups, receiving either anti-CD24 monoclonal antibody (clone SN3, Creative Diagnostics) or mouse IgG1 isotype control (clone MOPC-21, BioXcell). On day 5 after engraftment, mice received an initial dose of 200 μg and were subsequently treated every other day at a dose of 400 μg for two weeks. Bioluminescence imaging was performed throughout the study and after treatment withdrawal in order to assess tumour growth.

In vivo immunocompetent growth experiments. Female C57Bl/6 mice, 6–8 weeks of age were injected intraperitoneally with 1 × 10^6 ID8-WT-tdTomato-luc$^+$ or ID8-ΔCd24a-tdTomato-luc$^+$ cells in PBS. Tumour growth was measured by weekly bioluminescence imaging, beginning two weeks after engraftment.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

All primary data for all figures and supplementary figures are available from the corresponding authors upon request.

27. Goldman, M., Craft, B., Brooks, A. N., Zhu, J. & Haussler, D. The USCS Xena Platform for cancer genomics data visualization and interpretation. Preprint at https://doi.org/10.1101/326470v3 (2018).
28. Martinez, F. O. Analysis of gene expression and gene silencing in human macrophages. *Curr. Protoc. Immunol.* **96**, 14.28.1–14.28.23 (2012)
29. Brinkman, E. K., Chen, T., Amendola, M. & van Steensel, B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res.* **42**, e168 (2014).

Acknowledgements We thank the members of the Weissman laboratory, the Stanford Stem Cell Institute, R. L. Maute and K. S. Kao for advice and discussions; A. McCarty, T. Naik and L. Quinn for technical and logistical support; I. Wapnir for providing samples from patients with breast cancer; G. Wernig for providing human ascites samples; and G. Krampitz for the APL1 cell line. The research reported in this publication was supported by the Virginia D. K. Ludwig Fund for Cancer Research (NIHR01CA086017 and NIHGR01GM100315) and the NIH/NCI Outstanding Investigator Award (R35CA220434 to I.L.W.); the Stanford Medical Scientist Training Program (T32GM007365 to A.A.B); the National Cancer Institute (F30CA232472 to A.A.B); and the Program in Translational and Experimental Hematology T32

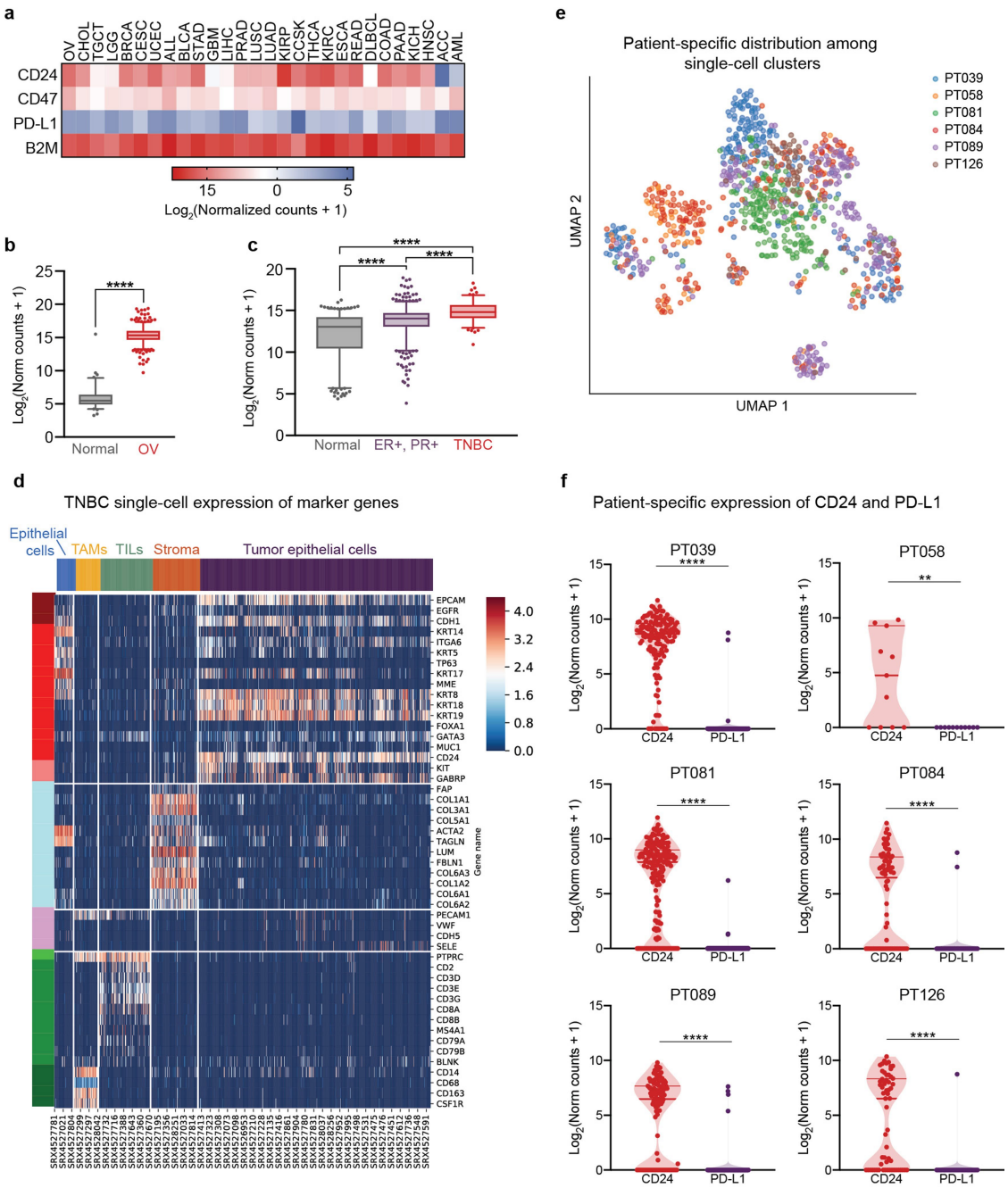
from the National Heart, Lung, and Blood Institute (1T32HL120824 to B.W.Z.). The contents of this manuscript are solely the responsibility of the authors.

Author contributions A.A.B. wrote the manuscript. A.A.B. and I.L.W. conceived and designed all experiments. A.A.B., R.E.B. and M.M. performed all flow cytometry analyses, generated human macrophages, and performed in vitro phagocytosis assays. A.A.B., R.E.B. and M.M. performed in vivo experiments and S.A.B. assisted with in vivo phagocytosis measurements. B.W.Z. assisted with the design of neuraminidase experiments. M.K. performed single-cell RNA-sequencing analysis. J.H. assisted with live-cell microscopy. L.J.B. assisted with statistical analysis and manuscript preparation. O.D. and V.K. provided primary human ovarian cancer samples. I.L.W. supervised the research and edited the manuscript.

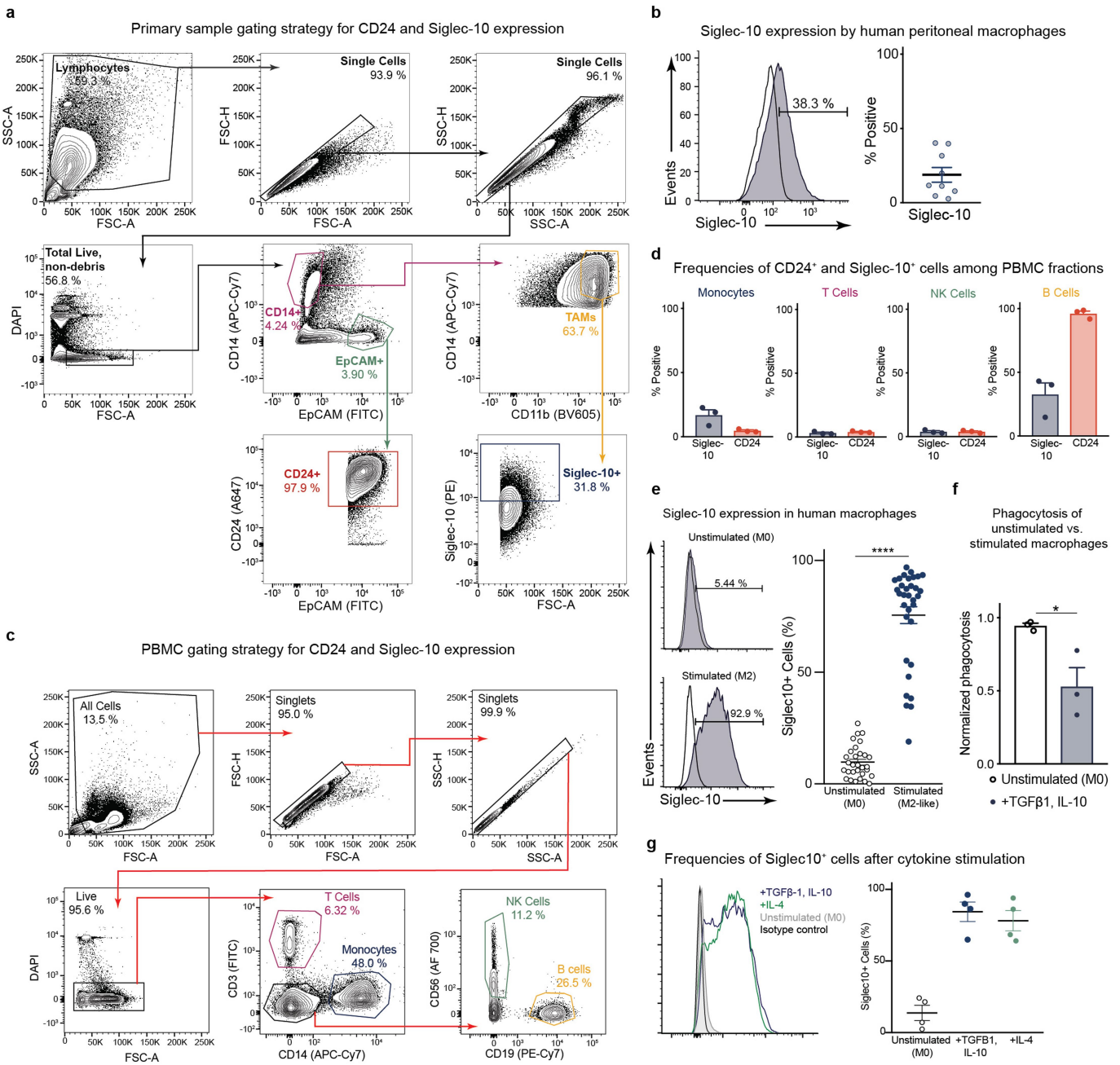
Competing interests A.A.B. and I.L.W. are co-inventors on a patent application (62/684,407) related to this work. I.L.W. is a founder, director, stockholder and consultant of Forty Seven, a cancer immunotherapy company.

Additional information

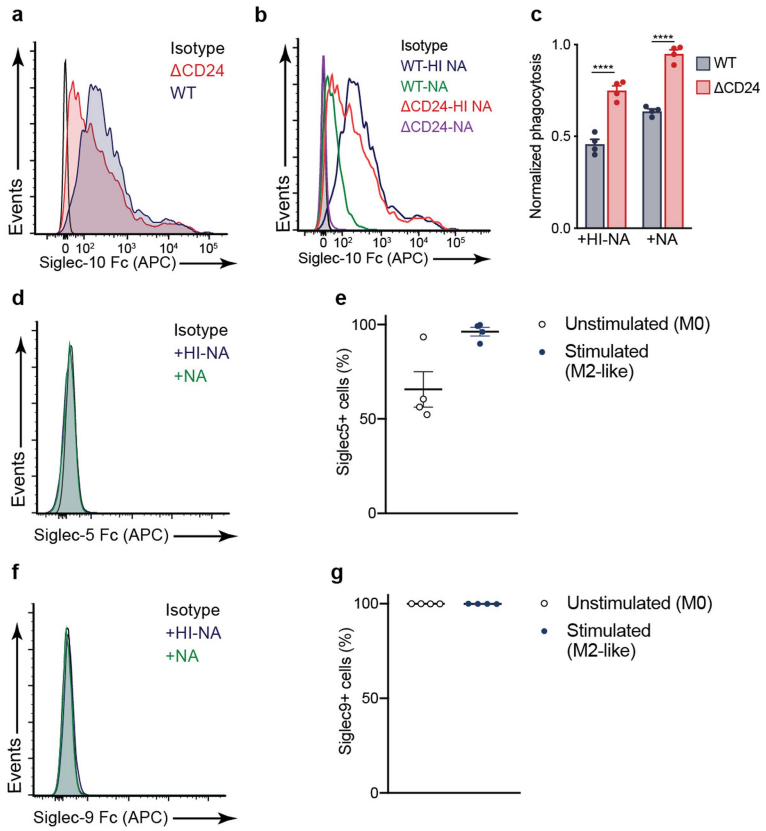
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1456-0.
Correspondence and requests for materials should be addressed to I.L.W.
Peer review information *Nature* thanks Gregory Beatty, Heinz Läubli and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



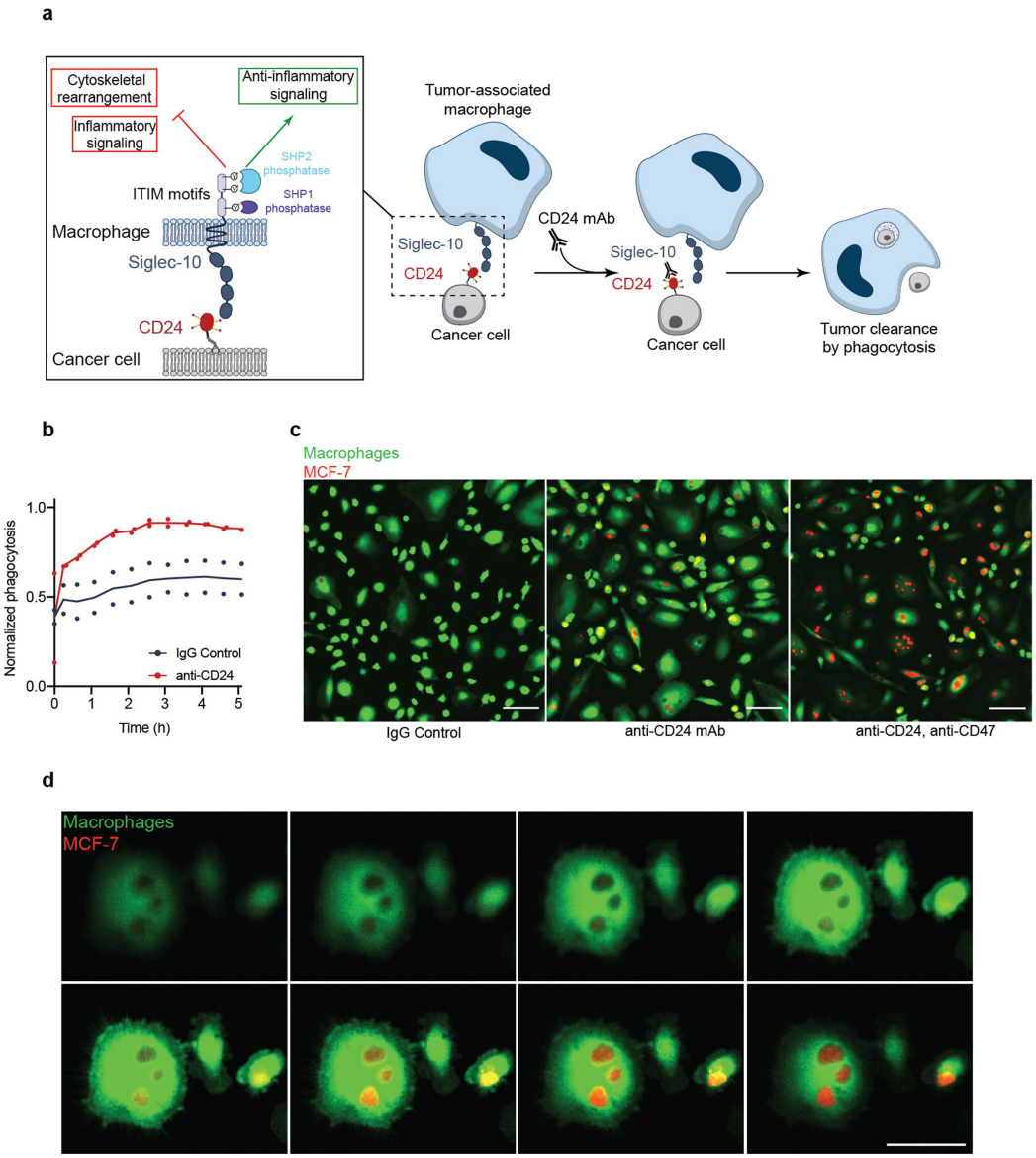
Extended Data Fig. 1 | Expression of innate immune checkpoints in human cancer. a, Heat map of expression ($\log_2$(normalized counts + 1)) of *CD24* from bulk TCGA and TARGET studies, as compared to known innate immune checkpoint molecules *CD47*, *PD-L1* and *B2M* (tumour study abbreviations and *n* values are defined in Supplementary Table 1). **b**, Expression levels of *CD24* in ovarian cancer (OV, red box plot, $n = 419$) in comparison with ovarian tissue from healthy individuals (grey box plot, $n = 89$), boxes show the median and whiskers indicate the 95th and 5th percentiles, $^{****}P < 0.0001$, unpaired, two-tailed Student's *t*-test. **c**, Expression levels of *CD24* in TNBC (red box plot, $n = 124$) in comparison with ER^+PR^+ breast cancer (purple box plot, $n = 508$) and healthy breast cells (grey box plot, $n = 293$). Each symbol represents an individual patient sample, boxes show the median and whiskers indicate the 95th and 5th percentiles, $^{****}P < 0.0001$, one-way ANOVA with multiple comparisons correction, $F_{(2,922)} = 95.80$. **d**, Heat map of marker gene expression (*y* axis) across TNBC single cells (*x* axis) and cell clusters identified (top). **e**, UMAP dimension 1 and 2 plots displaying all TNBC cells analysed from six patients ($n = 1{,}001$ single cells); cell clusters are coloured by cell patient (for key, see right). **f**, *CD24* compared with *PD-L1* expression in the 'Tumour epithelial cell' cluster for individual TNBC patients. Number of single cells analysed: PT039, $n = 151$ cells; PT058, $n = 11$ cells; PT081, $n = 196$ cells; PT084, $n = 84$ cells; PT089, $n = 117$ cells; PT126, $n = 60$ cells. $^{**}P < 0.01$, $^{****}P < 0.0001$. Data are violin plots showing median expression ($\log_2$(normalized counts +1) and quartiles (paired, two-tailed *t*-test).



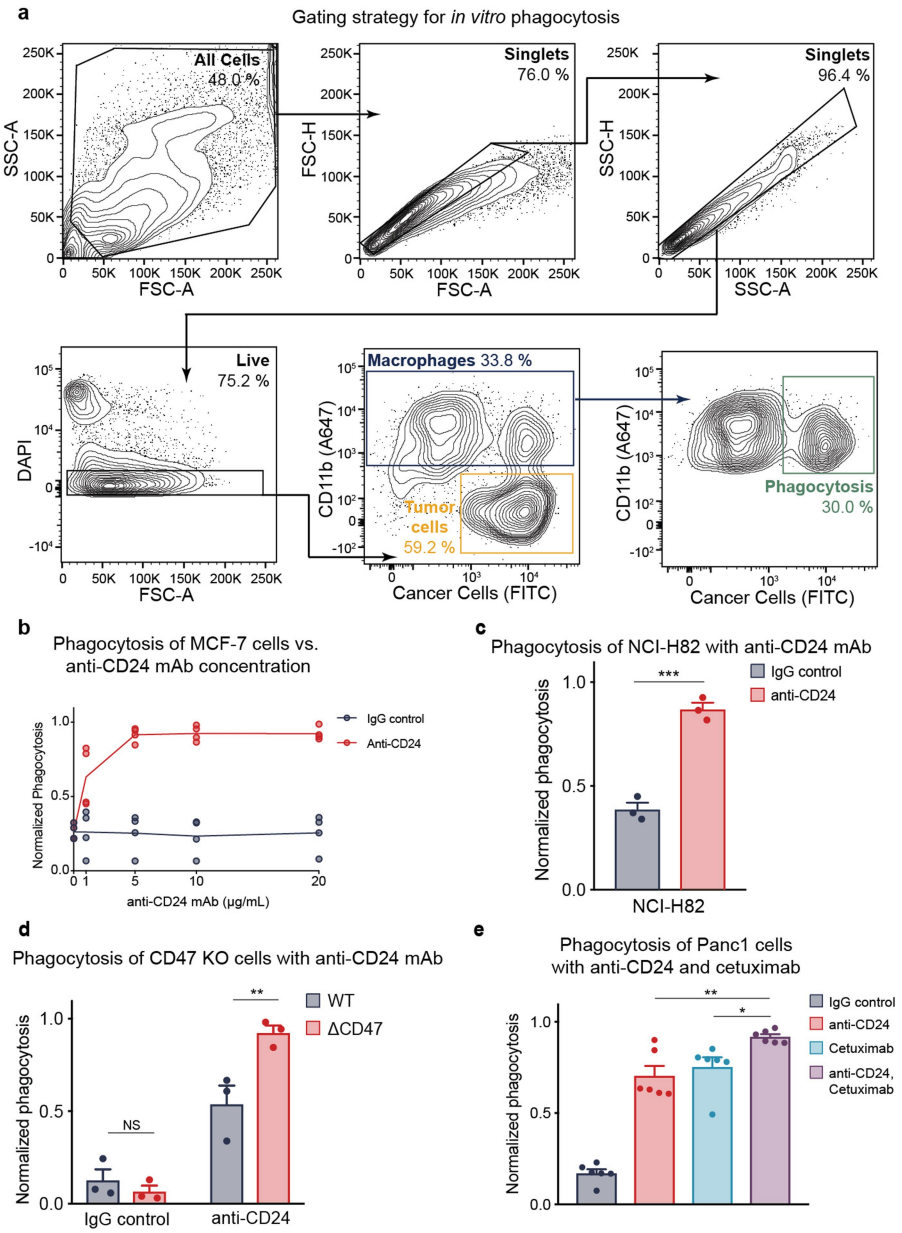
Extended Data Fig. 2 | Flow-cytometry analysis of CD24 and Siglec-10 expression in human tumours and primary immune cells. **a**, Gating strategy for $CD24^+$ cancer cells and Siglec-10^+ TAMs in primary human tumours; after debris and doublet removal, cancer cells were assessed as $DAPI^-CD14^-EpCAM^+$ and TAMs were assessed as $DAPI^-EpCAM^-CD14^+CD11b^+$. Plots are representative of six experimental replicates. **b**, Left, representative flow-cytometry histogram measuring the expression of Siglec-10 (blue shaded curves) versus isotype control (black lines) by non-cancerous peritoneal macrophages; numbers above bracketed line indicate the percentage of macrophages positive for expression of Siglec-10. Right, frequency of peritoneal macrophages positive for Siglec-10 among all peritoneal macrophages as defined by isotype controls ($n = 9$ donors). **c**, Gating strategy for $CD24^+$ cells and Siglec-10^+ cells among PBMC cell types; after debris and doublet removal, monocytes were assessed as $DAPI^-CD3^-CD14^+$; T cells were assessed as $DAPI^-CD14^-CD3^+$; natural killer (NK) cells were assessed as $DAPI^-CD14^-CD3^-CD56^+$; B cells were assessed as $DAPI^-CD56^-CD14^-CD3^-CD19^+$. Plots are representative of two experimental replicates. **d**, Frequency of PBMC cell types positive for Siglec-10 (blue shaded bars) or CD24 (red shaded bars) out of total cell type ($n = 3$ donors). **e**, Left, flow-cytometry-based measurement of the surface expression of Siglec-10 on primary human donor-derived macrophages either unstimulated (top) or after stimulation with M2-polarizing cytokines TGFβ1 and IL-10 (bottom), numbers above bracketed line indicate the per cent of $CD11b^+$ macrophages positive for expression of Siglec-10. Right, frequency of primary human donor-derived macrophages positive for Siglec-10 either without stimulation (unstimulated, M0) or following stimulation with TGFβ1 and IL-10 (stimulated, M2-like) ($n = 30$ unstimulated donors, 33 stimulated donors; unpaired, two-tailed Student's t-test, $^{****}P < 0.0001$, data are mean ± s.e.m.). **f**, Flow-cytometry-based measurement of phagocytosis of MCF-7 cells by unstimulated donor-derived macrophages (white data points) versus TGFβ1 and IL-10-stimulated donor-derived macrophages ($n = 3$ donors, unpaired, one-tailed t-test, $^*P = 0.0168$). **g**, Left, flow-cytometry-based measurement of the surface expression of Siglec-10 on matched, primary donor-derived macrophages either unstimulated (grey shaded curve), or after stimulation with TGFβ1 and IL-10 (blue line), or IL-4 (green line). Right, frequency of matched, human donor-derived macrophages positive for Siglec-10 either without stimulation (unstimulated, M0), or after stimulation with TGFβ1 and IL-10 (blue dots), or stimulated with IL-4 ($n = 4$ donors). Data are mean ± s.e.m.



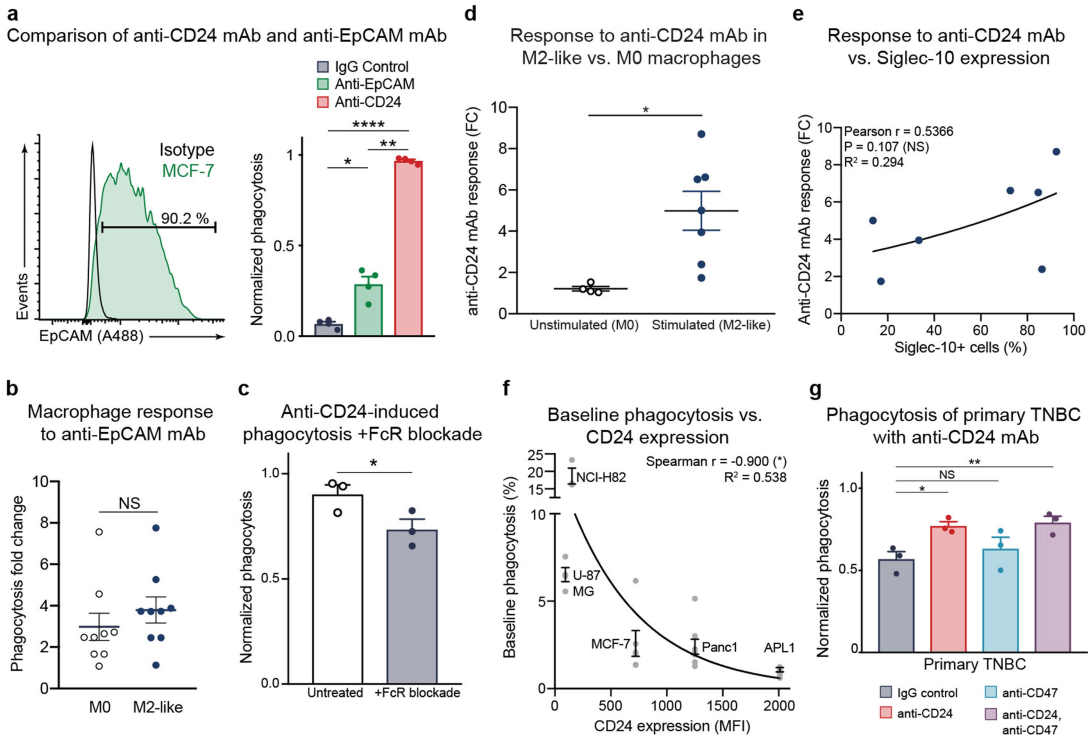
Extended Data Fig. 3 | Siglec-10 binds to CD24 expressed on MCF-7 cells. a, Flow cytometry histogram measuring the binding of Siglec-10 to wild-type MCF-7 cells (blue shaded curve) versus MCF-7(ΔCD24) cells (red shaded curve). Data are representative of two experimental replicates. **b**, Merged flow cytometry histogram measuring the binding of Siglec-10–Fc to wild-type MCF-7 cells treated with heat-inactivated neuraminidase (WT-HI NA, blue line), wild-type MCF-7 cells treated with neuraminidase (WT-NA, green line), MCF-7(ΔCD24) cells treated with heat-inactivated neuraminidase (red line, ΔCD24-HI NA), and MCF-7(ΔCD24) cells treated with neuraminidase (purple line, ΔCD24-NA) as compared to isotype control (black line). Data are representative of two experimental replicates. **c**, Flow-cytometry-based measurement of phagocytosis of $CD24^+$ parental MCF-7 cells (WT) and $CD24^-$ (ΔCD24) MCF-7 cells by co-cultured human macrophages in the presence of neuraminidase (+NA) or heat-inactivated neuraminidase (+HI-NA) ($n = 4$ donors; two-way ANOVA with multiple comparison's correction, cell line $F_{(1,12)} = 180.5$, treatment $F_{(1,12)} = 71.12$, $^{****}P < 0.0001$, data are mean ± s.e.m.). **d**, **f**, Representative flow cytometry histogram measuring the binding of Siglec-5 (**d**) or Siglec-9 (**f**) to wild-type MCF-7 cells treated with either vehicle (blue shaded curve) or neuraminidase (green shaded curve). Data are representative of two experimental replicates. **e**, **g**, Frequency of macrophages positive for Siglec-5 (**e**) or Siglec-9 (**g**) among unstimulated M0 macrophages or stimulated M2-like macrophages ($n = 4$ donors). Data are mean ± s.e.m.



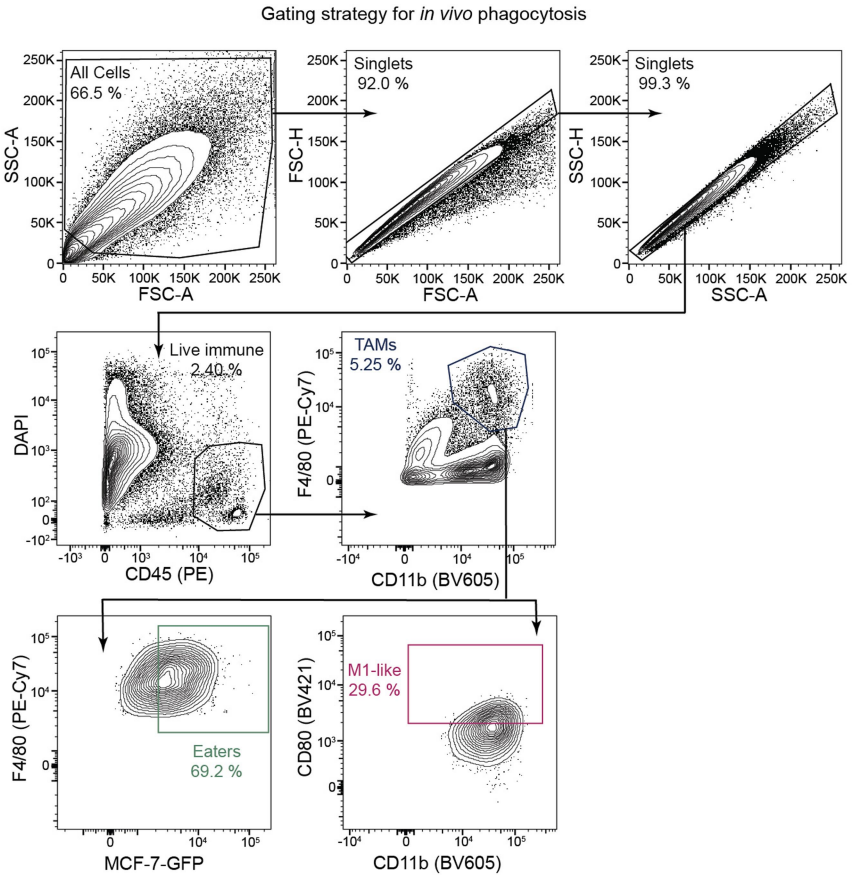
Extended Data Fig. 4 | Anti-CD24 monoclonal antibodies promote phagocytic clearance of cancer cells over time. a, Schematic of the inhibition of phagocytosis by CD24–Siglec-10. The inhibitory receptor Siglec-10 engages its ligand CD24 on cancer cells, leading to phosphorylation of the two immunoreceptor tyrosine-based inhibition motifs in the cytoplasmic domain of Siglec-10 and subsequent anti-inflammatory, anti-phagocytic signalling cascades mediated by SHP-1 and SHP-2 phosphatases; upon the addition of a CD24 blocking antibody, macrophages are disinhibited and are thus capable of phagocytosis-mediated tumour clearance. **b**, Quantification of phagocytosis events of MCF-7 cells treated with anti-CD24 mAb (red curve) versus IgG control (blue curve) as measured by live-cell microscopy over time, normalized to maximum measured phagocytosis events per donor (n = two donors; P value computed by two-way ANOVA of biological replicates, $F_{(1,24)} = 65.02$). Line is the mean of two biological replicates with individual replicates shown. **c**, Representative fluorescence microscopy images of in vitro phagocytosis of MCF-7 cells (mCherry$^+$, red) by macrophages (Calcein, AM; green) in the presence of IgG control (left), anti-CD24 mAb (middle), or anti-CD24 mAb and anti-CD47 mAb (right), after 6 h of co-culture. Experiment was repeated with three donors. Scale bar, 100 μm. **d**, Representative Z-stack images collected from high-resolution confocal fluorescence microscopy of macrophage phagocytosis demonstrating engulfment of whole MCF-7 cells (mCherry$^+$, red) by macrophages (Calcein, AM; green). Experiment was repeated with three donors. Scale bar, 50 μm.



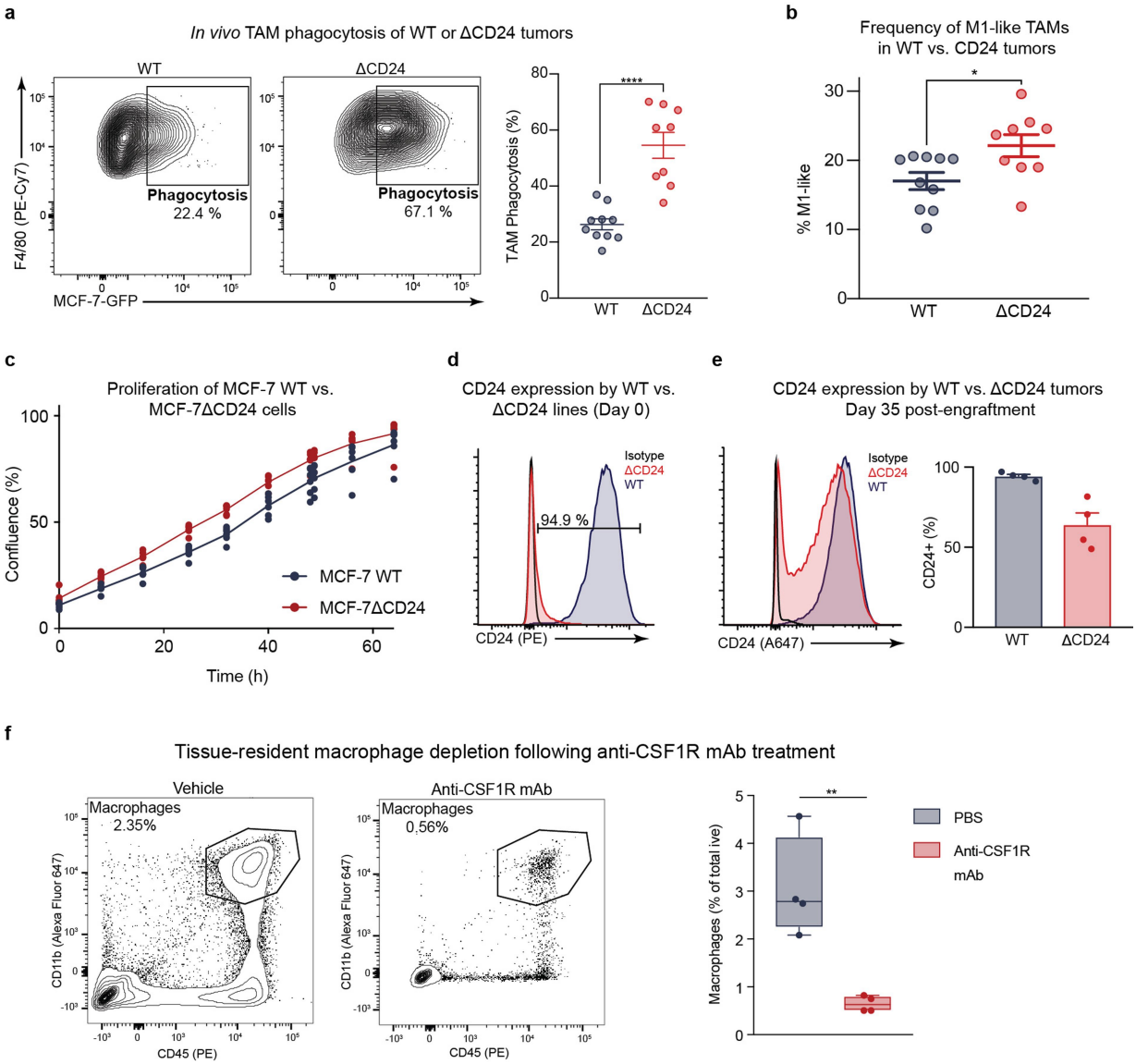
Extended Data Fig. 5 | CD24 antibody blockade of CD24–Siglec-10 signalling promotes dose-responsive enhancement of phagocytosis. **a**, Gating strategy for in vitro phagocytosis assay. Following debris and doublet removal, phagocytosis was assessed as the frequency of $DAPI^{-}CD11b^{+}FITC^{+}$ events out of all $DAPI^{-}CD11b^{+}$ events. Numbers indicate frequency of events out of previous gate. Plots are representative of at least 10 experimental replicates. **b**, Dose–response relationship of anti-CD24 mAb on phagocytosis of MCF-7 cells, concentrations listed on the *x* axis as compared to IgG control ($n = 3$ donors). Connecting line is mean. **c**, Flow-cytometry-based measurement of phagocytosis of NCI-H82 cells by donor-derived macrophages ($n = 3$ donors) in the presence of anti-CD24 mAb as compared to IgG control; each symbol represents an individual donor (paired, two-tailed Student's *t*-test, $^{***}P = 0.0001$). Data are mean ± s.e.m. **d**, Flow-cytometry-based measurement of phagocytosis of $CD24^{+}$ parental MCF-7 cells (WT) and $CD47^{-}$ (ΔCD47) MCF-7 cells by co-cultured human macrophages, in the presence or absence of anti-CD24 mAb (horizontal axis), ($n = 4$ donors; two-way ANOVA with multiple comparisons correction, cell line $F_{(1,8)} = 6.490$; treatment $F_{(1,8)} = 98.73$, $^{**}P = 0.0054$). Data are mean ± s.e.m. **e**, Flow-cytometry-based measurement of phagocytosis of Panc1 pancreatic adenocarcinoma cells in the presence of anti-CD24 mAb, cetuximab (anti-EGFR), or both anti-CD24 mAb and cetuximab, as compared to IgG control ($n = 6$ donors) (one-way ANOVA with multiple comparisons correction, $F_{(3,20)} = 66.10$, $^{*}P = 0.0373$, $^{**}P = 0.0057$, data are mean ± s.e.m.).



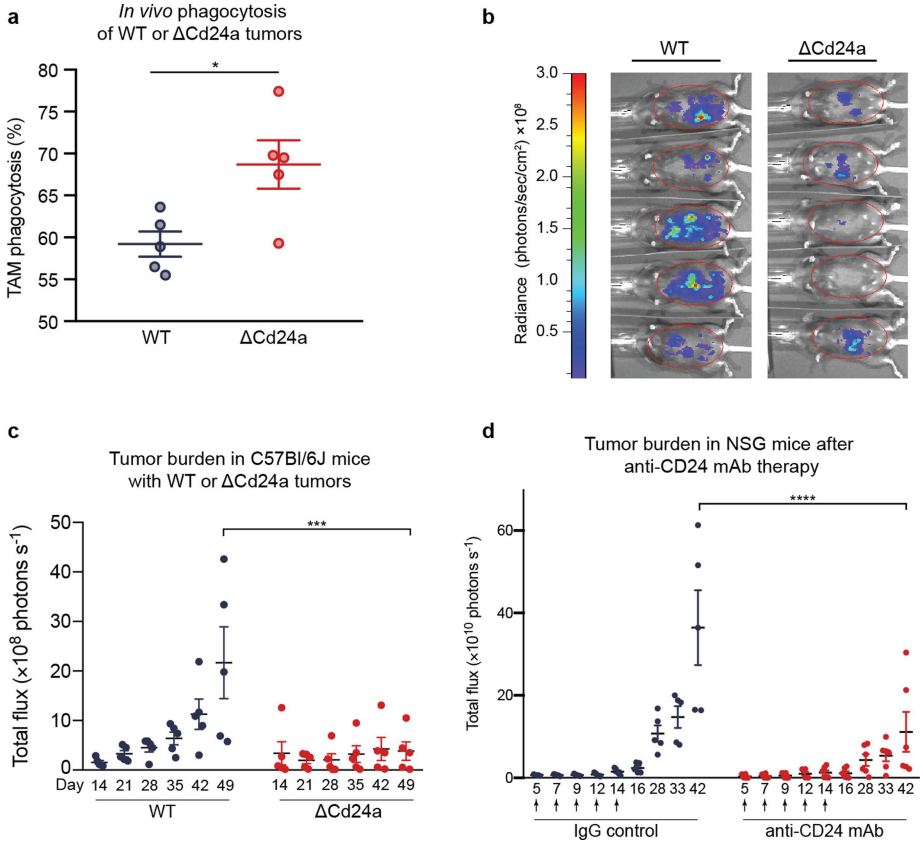
Extended Data Fig. 6 | The opsonization effect of anti-CD24 mAb is minor and CD24 blockade promotes phagocytosis of primary TNBC. a, Left, representative flow cytometry histogram measuring the expression of EpCAM (green shaded curve) by parental MCF-7 cells; number above bracketed line indicates the percentage of MCF-7 cells positive for expression of EpCAM. Right, flow cytometry–based measurement of phagocytosis of parental MCF-7 cells by co-cultured human macrophages, in the presence of either IgG control, anti-EpCAM mAb or anti-CD24 mAb ($n = 4$ donors; repeated measures ANOVA with multiple comparisons correction, $F_{(2,9)} = 340.9$, $^{*}P = 0.0287$, $^{**}P = 0.0015$, $^{****}P < 0.0001$). Data are mean ± s.e.m. **b**, Fold change in phagocytosis by M0 (unstimulated) or M2-like (TGFβ1, IL-10-stimulated) macrophages upon the addition of anti-EpCAM mAb as compared to IgG control ($n = 9$ donors, paired, two-tailed t-test). Data are mean ± s.e.m. **c**, Flow-cytometry-based measurement of anti-CD24 mAb-induced phagocytosis of MCF-7 cells by untreated macrophages (white bar) versus macrophages treated with anti-CD16/32 mAb (+FcR blockade, blue bar) ($n = 3$ macrophage donors. Paired, two-tailed t-test. Each point represents an individual donor. $^{*}P = 0.0358$). Data are mean ± s.e.m. **d**, Response to anti-CD24 mAb by M2-like macrophages compared with M0 macrophages; each symbol represents an individual donor ($n = 4$, M0 donors; $n = 6$, M2-like donors; unpaired, two-tailed Student's t-test, $^{*}P = 0.0160$). **e**, Pearson correlation between stimulated (M2-like) donor-derived macrophage Siglec-10 expression and response to anti-CD24 mAb as computed by the phagocytosis fold change between anti-CD24 mAb treatment and IgG control ($n = 7$ donors); exponential growth curve is shown. **f**, Spearman correlation between cancer cell CD24 expression and baseline, un-normalized phagocytosis levels (IgG control) averaged across all donors per cell line. Exponential growth equation is shown (n values are the same as in Fig. 3b and Extended Data Fig. 5c, $^{*}P = 0.0417$). Data are mean ± s.e.m. **g**, Flow-cytometry-based measurement of phagocytosis of a sample of primary TNBC cells from a patient, in the presence of anti-CD24 mAb, anti-CD47 mAb, or both anti-CD24 mAb and anti-CD47 mAb, as compared to IgG control ($n = 3$ macrophage donors challenged with $n = 1$ primary TNBC donor; repeated-measures one-way ANOVA with multiple comparisons correction, $F_{(1.217,2.434)} = 26.17$). Each point represents an individual donor. $^{*}P = 0.0434$, $^{**}P = 0.0028$. Data are mean ± s.e.m.



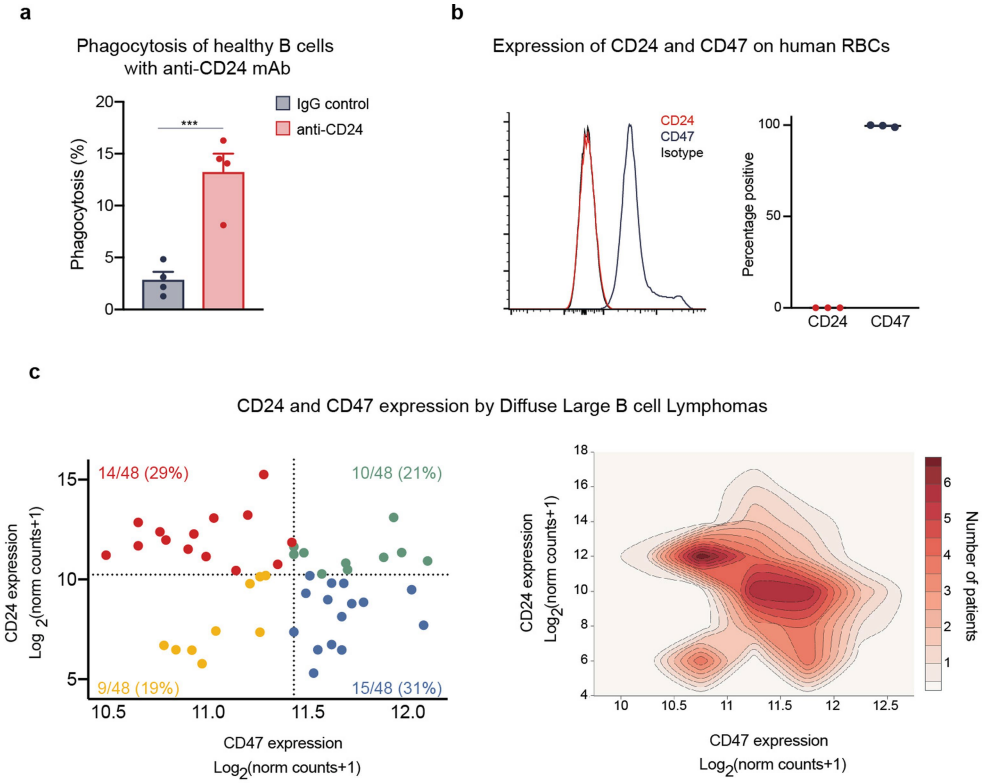
Extended Data Fig. 7 | Gating strategy for in vivo phagocytosis. Gating strategy for in vivo TAM phagocytosis of MCF-7 cells; after debris and doublet removal, TAM phagocytosis is assessed as the frequency of $DAPI^{-}CD11b^{+}F4/80^{+}GFP^{+}$ events out of total $DAPI^{-}CD11b^{+}F4/80^{+}$ events; M1-like TAMs assessed as $DAPI^{-}CD11b^{+}F4/80^{+}CD80^{+}$, numbers indicate frequency of events out of the previous gate. Plots are representative of three experimental replicates.



Extended Data Fig. 8 | Characterization of MCF-7 wild-type and MCF-7(ΔCD24) cells in vitro and in vivo. a, Representative flow cytometry plots demonstrating TAM phagocytosis in GFP-luciferase^{+}CD24^{+} (WT) MCF-7 tumours (left) versus CD24^{-} (ΔCD24) MCF-7 tumours (middle), numbers indicate frequency of phagocytosis events out of all TAMs. Right, frequency of phagocytosis events out of all TAMs in wild-type tumours versus ΔCD24 tumours 28 days after engraftment (WT, $n = 10$; ΔCD24, $n = 9$; unpaired, two-tailed Student's t-test, $^{****}P < 0.0001$). **b**, Frequency of TAMs positive for CD80 (M1-like) as per gating in **a**, among all TAMs macrophages as defined by fluorescence minus one controls (WT, $n = 10$; ΔCD24, $n = 9$; unpaired, two-tailed Student's t-test, $^{*}P < 0.0203$). Data are mean ± s.e.m. **c**, In vitro proliferation rates of MCF-7 wild-type and MCF-7(ΔCD24) as assessed by a plot of confluence percentage over time ($n = 6$ technical replicates, one experimental replicate). Individual technical replicates are shown, the connecting line indicates the mean. **d**, Flow-cytometry-based measurement of the surface expression of CD24 on MCF-7 cells (blue shaded curve) versus CD24 knockout cells (ΔCD24) (red shaded curve) before tumour engraftment as compared to isotype control (black line), numbers above the bracketed line indicate the percentage of MCF-7 wild-type cells positive for expression of CD24. The plot is representative of ten experimental replicates. **e**, Left, representative flow-cytometry histogram of the surface expression of CD24 on day-35 wild-type MCF-7 tumours (blue shaded curve) versus day-35 CD24 knockout tumours (ΔCD24) (red shaded curve) as compared to isotype control (black line). Right, flow-cytometry-based measurement of the frequency of CD24^{+} cells among all cancer cells in day-35 wild-type tumours versus day 35 ΔCD24 tumours (WT, $n = 4$; ΔCD24, $n = 4$). Data are mean ± s.e.m. **f**, Representative flow cytometry plots of tissue-resident macrophages out of total live cells in vehicle-treated mice (left) compared with anti-CSF1R-treated mice (middle); numbers indicate frequency of CD11b^{+}F4/80^{+} macrophage events out of total live events. Right, frequency of TAMs (CD11b^{+}F4/80^{+}) out of total live cells in vehicle-treated mice ($n = 5$, blue shaded box plot) versus anti-CSF1R-treated mice ($n = 4$, red shaded box plot) as measured by flow cytometry. $^{**}P < 0.01$. Box plots depict mean and range.



Extended Data Fig. 9 | Validation of CD24 inhibition in in vivo models of ovarian and breast cancer. a, In vivo phagocytosis of wild-type or ΔCd24a cancer cells by mouse TAMs. Flow cytometry–based measurement of in vivo phagocytosis of $CD24^{+}GFP^{+}$ ID8 cells (WT) versus $CD24^{-}GFP^{+}$ ID8 cells (ΔCd24a) by mouse peritoneal macrophages ($n = 5$ mice; unpaired, two-tailed Student's *t*-test with multiple comparisons correction, $^{*}P = 0.0196$). **b**, Representative bioluminescence image of tumour burden in C57Bl/6 mice with ID8 wild-type versus ID8(ΔCd24a) tumours (image taken 49 days after engraftment and representative of one experimental replicate). **c**, Burden of ID8 wild-type tumours (blue) versus ID8(ΔCd24a) tumours (red) as measured by bioluminescence imaging (WT, $n = 5$; ΔCd24a, $n = 5$; two-way ANOVA with multiple comparisons correction, tumour genotype $F_{(1,48)} = 10.70$, $^{***}P = 0.0001$). Data are representative of one experimental replicate. **d**, Extended measurement (as in Fig. 4e) of burden of MCF-7 wild-type tumours treated with IgG control (blue) versus anti-CD24 mAb (red) as measured by bioluminescence (IgG control, $n = 5$; anti-CD24 mAb, $n = 5$; days on which treatment was administered are indicated by arrows below the *x* axis; data are of one experimental cohort; two-way ANOVA with multiple comparisons correction, tumour treatment $F_{(1,81)} = 16.75$). $^{****}P < 0.0001$. Data are mean ± s.e.m.



Extended Data Fig. 10 | Anti-CD24 mAb induces B cell clearance but does not bind human red blood cells, and CD47 and CD24 demonstrate inversely correlated expression in human diffuse large B-cell lymphoma. a, Flow-cytometry-based measurement of phagocytosis of B cells ($n = 4$ donors, pooled) by donor-derived macrophages ($n = 4$ donors) in the presence of anti-CD24 mAb as compared to IgG control; each symbol represents an individual donor (paired, two-tailed Student's *t*-test, $^{***}P = 0.0008$). **b**, Left, representative flow cytometry histogram measuring the expression of CD24 (red line) and CD47 (blue line) by human red blood cells (RBCs); right, flow-cytometry-based measurement of the frequency of $CD24^+$ compared with $CD47^+$ RBCs out of total RBCs ($n = 3$ donors). Data are mean ± s.e.m. **c**, Left, expression levels in $\log_2$(normalized counts + 1) of *CD24* and *CD47* in diffuse large B cell lymphomas from TCGA ($n = 48$); data are divided into quadrants by median expression of each gene as indicated by dotted lines. The number and percentage of total patients in each quadrant are indicated on the plot. Each dot indicates a single patient. Right, two-dimensional contour plot of expression levels of *CD24* and *CD47* in the large B cell lymphoma samples featured in the left plot.

nature research

Corresponding author(s): Irving L. Weissman

Last updated by author(s): Jun 13, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist .

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Flow cytometry data was collected using FACSDiva 8.0.1 (BD). Bioluminescence data was collected using Living Image 4.2.
Data analysis	Flow cytometry data was analyzed using FACSDiva 8.0.1 (BD) and FlowJo 10 (Tree Star). Bioluminescence data was analyzed using Living Image 4.2. All graphs were generated and analyzed using GraphPad Prism 8. Indel analysis for gene knockouts was performed using TIDE 2.0.1. Single-cell RNA sequencing analysis was performed with STAR 2.5.3a. Contour plots were generated using Plotly 2018.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All primary data for all figures and supplementary figures are available from the corresponding authors upon request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes were modeled after those from existing publications regarding in vitro immune killing assays and in vivo tumor growth assays, and an independent statistical method was not used to determine sample size. In our experience with in vitro measurements of phagocytosis, we have found that assaying human macrophages from 3 donors is sufficient for studies of antibody efficacy based off of observed variability among donors.
Data exclusions	As listed in the Methods, phagocytosis assays were performed in a minimum of technical triplicate for a minimum of 3 human donors per treatment group. In some cases, donors or specific technical replicates were excluded on the pre-established criterion that they were found to be a significant outlier by the GraphPad Outlier Calculator (https://www.graphpad.com/quickcalcs/Grubbs1.cfm). In some cases, additional replicates of specific phagocytosis assay conditions were repeated as part of pilot experiments, or as confirmatory replicates, but only a discrete set of data performed under identical conditions was specifically reported. For in vivo experiments, individual mice were removed from the study either prior to treatment, if found to be an engraftment outlier by bioluminescence imaging, or from the final analysis if, at end point, the mouse was found to be a significant outlier with regards to tumor growth. These exclusion criteria were established prior to tumor engraftment. All outlier calculations for in vivo experiments were performed using the GraphPad Outlier Calculator (https://www.graphpad.com/quickcalcs/Grubbs1.cfm). In some cases across additional experiments, including pilot experiments, additional mice were engrafted subcutaneously with relevant cell lines and followed for non-standard periods of time, or assessed for tumor growth at non-standard intervals, but only a discrete set of mice assessed under identical conditions was reported.
Replication	In vitro phagocytosis assays were performed in technical triplicate for a minimum of 3 human donors per treatment group with similar results and responses observed across donors and replicates. In vitro phagocytosis assays were performed across multiple experimental replicates, when possible, with the exceptions of the phagocytosis assays shown in Figure 2d (4 biological replicates, one experimental replicate), Figure 2g (4 biological replicates, one experimental replicate), Figure 2b (U-87 only; 3 biological replicates, one experimental replicate), Extended Data Figure 2e (3 biological replicates, one experimental replicate), Extended Data Figure 3c (4 biological replicates, one experimental replicate), Extended Data Figure 5c (3 biological replicates, one experimental replicate), Extended Data Figure 5d,f (4 biological replicates, one experimental replicate), Extended Data Figure 9a (4 biological replicates, one experimental replicate). Staining and recombinant Siglec binding experiments were performed in at least 2 experimental replicates. Automated live cell microscopy experiments were performed across at least technical and biological duplicates. Whenever practical for in vivo experiments, multiple cohorts across experimental replicates were performed. The number of cohorts performed is listed in the figure legends pertinent for each in vivo experiment. We observed similar results across cohorts and across individual mice within each cohort, as represented in the figures.
Randomization	For macrophage depletion experiments, mice pre-treated with either vehicle or anti-CSF1R mAb were randomized amongst treatment cohorts prior to engraftment with WT or CD24 KO MCF-7 tumors. Similarly, mice engrafted with MCF-7 tumors were randomized prior to treatment with anti-human CD24 mAb.
Blinding	All experiments, including in vivo experiments, were performed by unblinded investigators as all experiments in this work contained internal controls to allow for quantification and data analysis.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
[ ]	[x] Antibodies
[ ]	[x] Eukaryotic cell lines
[x]	[ ] Palaeontology
[ ]	[x] Animals and other organisms
[ ]	[x] Human research participants
[x]	[ ] Clinical data

Methods

n/a	Involved in the study
[x]	[ ] ChIP-seq
[ ]	[x] Flow cytometry
[x]	[ ] MRI-based neuroimaging

Antibodies

Antibodies used	All antibodies used in this work, clone, application, and supplier are listed in Supplementary Table 1.
Validation	The anti-human CD24 antibody (Clone SN3, Novus Bio (NB100-64861) and Creative Biolabs (CSC-S170)) used for staining and treatment studies in this work was validated by Novus Bio in human peripheral blood granulocytes. This antibody was also validated by staining unmodified MCF-7 cells versus CD24 knockout MCF-7 cells (dilution assessed in this work 1:50). The SN3

antibody was confirmed to not bind to mouse CD24a-expressing ID8 cells by flow cytometry. The CD24a antibody (Clone M1/69, BioLegend (101814)) was validated by staining unmodified ID8 cells versus CD24a knockout ID8 cells (dilution assessed in this work 1:100). The anti-human CD47 antibody used for treatments (Clone 5F9-G4, in house) is a clinical trial-grade humanized antibody which was validated as described in Liu et al. nature research | reporting summary October 2018 PLoS One (2015). The anti-human CD47 antibody used for staining (Clone B6H12, eBioscience (17-0479-42)) was validated by Barkal et al. Nature Immunology (2018) by comparing staining (dilution assessed in this work 1:100) of unmodified versus CD47 knockout cells. The Siglec-10 antibody (Clone 5G6, Thermo Scientific (MA5-28236)) has been validated by Thermo Fisher Scientific by staining CHO cells modified to express human Siglec-10 (dilution assessed in this work 1:50). The anti-human CD45 antibody (Clone HI30, BioLegend (304008)), the anti-human CD56 antibody (Clone HCD56, BioLegend (318316)), the anti-human CD3 antibody (Clone UCHT1, BioLegend (300415)), and the anti-human CD19 antibody (Clone SJ25C1, BioLegend (363011)) were all validated by the manufacturer by staining human peripheral lymphocytes (dilution assessed in this work 1:100). The anti-human/mouse CD11b antibody (Clone M1/70, BioLegend (101220)) was validated by the manufacturer by staining C57BL/6 mouse bone marrow cells (dilution assessed in this work 1:100). The anti-human CD14 antibody (Clone M5E2, BioLegend (301819)) was validated by the manufacturer by staining human peripheral blood monocytes (dilution assessed in this work 1:100). The anti-human EpCAM antibody (Clone 9C4, BioLegend (324204)) and the anti-human EpCAM antibody (Clone VU-1D9, ThermoFisher Scientific (BMS171)) were validated by the manufacturer by staining the HT29 human colon carcinoma cell line (dilution assessed in this work 1:100). The anti-human Siglec-5 antibody (Clone 1A5, BioLegend (352003)) was validated by the manufacturer by staining human peripheral blood granulocytes. The anti-human Siglec-9 antibody (Clone K8, BioLegend (351503)) was validated by the manufacturer by staining human peripheral blood monocytes. The anti-mouse CD45 antibody (Clone 30-F11, BioLegend (103106)) was validated by the manufacturer by staining C57BL/6 mouse splenocytes (dilution assessed in this work 1:100). The anti-mouse CD80 antibody (Clone 16-10A1, BioLegend (104725)) and the anti-mouse F4/80 antibody (Clone BM8, BioLegend (123114)) were validated by the manufacturer by staining thioglycolate-induced Balb/c mouse peritoneal macrophages (dilution assessed in this work 1:100). The anti-mouse CSF1R antibody (Clone AFS98, BioXCell (BE0213)) was validated by the investigators through FACS measurements of the frequency of tissue resident macrophages after 18 days of IP treatment with CSF1R antibody as compared to vehicle-treated mice.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	All cell lines used in this work were obtained from ATCC, with the exception of the APL1 human pancreatic neuroendocrine tumor line which was derived from a primary patient tumor as described in Krampitz et al. PNAS (2016) and the ID8 murine ovarian carcinoma cell line which was a gift from the laboratory of O. Dorigo.
Authentication	Cell lines were not independently authenticated beyond the identity provided from ATCC. The APL1 cell line was not independently authenticated beyond that performed in Krampitz et al. PNAS (2016). The ID8 murine ovarian carcinoma cell line was not independently authenticated.
Mycoplasma contamination	Stocks of all cell lines were tested for mycoplasma contamination prior to submission. All were negative.
Commonly misidentified lines (See ICLAC register)	None of the cell lines used in this study are listed in the database of commonly misidentified cell lines.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Animals used in xenograft experiments were 6-10 week old females of the NOD-scid IL2rγ-null (NSG) background obtained from in house breeding stocks. Animals used for syngeneic experiments were 6-8 week old females of the C57BL/6 background obtained from the Jackson Laboratory.
Wild animals	This study did not involve wild animals.
Field-collected samples	This study did not involve samples collected in the field.
Ethics oversight	All experiments were carried out in accordance with ethical care guidelines set by the Stanford University Administrative Panel on Laboratory Animal Care. Specific protocol numbers available on request.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics	The primary human samples used in this work were all collected from female patients who had been diagnosed with ovarian cancer or breast cancer and who were operated on at Stanford University Medical Center. All patients were above 30 years of age and female. Information not protected by HIPAA (i.e. age, genotypic/molecular information) available on request.
Recruitment	Female patients with ovarian cancer and breast cancer identified by the surgeons (I. Wapnir, breast cancer; O. Dorigo, ovarian cancer; Human Immune Monitoring Center Biobank and Stanford Tissue Bank; breast cancer) were recruited for the IRB approved studies reported here.

Ethics oversight	The Human Immune Monitoring Center Biobank, the Stanford Tissue Bank, and Dr. Oliver Dorigo all received IRB approval from the Stanford University Administrative Panels on Human Subjects Research and complied with all ethical guidelines for human subjects research to obtain patient samples of ovarian cancer and breast cancer, and received informed consent from all patients. Specific IRB protocol numbers are available on request.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Please also see description of sample preparation included in the Methods. Briefly: FACS of primary human tumors/mouse tumors: Solid tumors were excised and mechanically dissociated using a straight razor prior to incubation with 10 mL RPMI supplemented with 10 mL of RPMI + 10 µg/mL DNaseI (Sigma Aldrich) + 25 µg/mL Liberase (Roche) for 30-60 min at 37°C. Single cell suspensions were blocked using species-matched anti-CD16/32 antibodies (TruStain fcX, BioLegend) for 10 minutes on ice prior to staining. Gates were set by fluorescence minus one controls for markers other than CD24 and Siglec-10 which were set based off of isotype controls. All samples were analyzed while in FACS buffer containing 1 µg/mL DAPI in order to exclude dead cells. Channel compensations were performed using single-stained UltraComp eBeads (Affymetrix) or cells. In vitro phagocytosis assay: All phagocytosis assay wells were stained with anti-human/mouse CD11b antibody (Clone M1/70, BioLegend) for 30 minutes on ice, prior to analysis. All samples were analyzed while in FACS buffer containing 1 µg/mL DAPI in order to exclude dead cells. Gates were set based off of fluorescence minus one controls. Data was analyzed using FlowJo (Treestar) and outliers among technical replicates in each treatment group were removed using GraphPad Outlier Calculator (http://graphpad.com/quickcalcs/Grubbs1.cfm).
Instrument	All samples were analyzed on an LSR Fortessa (BD) or Aria II SORP (BD).
Software	FACS data was collected using FACS Diva (BD) and analyzed using FACS Diva (BD) or FlowJo (Tree Star). Statistical analyses and plots were generated using GraphPad Prism 8.
Cell population abundance	The CD24-null MCF7-7 and ID8 CD24a-null cell lines were sorted based off of positive controls (WT versions of each cell line). After the initial knockout was performed, the CD24-null cell lines approximately 10% of the population, while in successive purification sorts, the CD24-null population was >70% of the population. No other populations were sorted for this manuscript.
Gating strategy	All gating strategies used in this work are included in the Extended Data Figures. Briefly: FACS of primary human tumors: The frequency of CD24+ cancer cells was measured as the number of DAPI–CD14–EpCAM+CD24+ cells out of total DAPI–CD14–EpCAM+ cells as determined by isotype controls. The frequency of Siglec-10+ TAMs was measured as the number of DAPI–EpCAM–CD14+CD11b+Siglec-10+ cells out of total DAPI–EpCAM–CD14+CD11b+ cells, as determined by isotype controls (Gating Strategy Extended Data Figure 2a). In vitro phagocytosis assays: Phagocytosis was defined as the frequency of the DAPI-CD11b+GFP+ population among all DAPI-CD11b+ cells (Gating Strategy Extended Data Figure 3). In vivo TAM phagocytosis assays: In vivo phagocytosis was defined as the DAPI-CD45+ CD11b+F480+GFP+ population out of total TAMs defined as DAPI-CD45+ CD11b+F480+ cells (Gating Strategy Extended Data Figure 5a). M1-like mouse TAMs: The frequency of M1-like mouse TAMs was measured as DAPI–, CD45+, CD11b+, F480+, CD80+ TAMs out of total TAMs defined as DAPI-CD45+ CD11b+F480+, as defined by fluorescence minus one controls (Gating Strategy Extended Data Figure 5a).

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

LETTER

https://doi.org/10.1038/s41586-019-1437-3

Dietary methionine influences therapy in mouse cancer models and alters human metabolism

Xia Gao[1], Sydney M. Sanderson[1,8], Ziwei Dai[1,8], Michael A. Reid[1], Daniel E. Cooper[2], Min Lu[3,4], John P. Richie Jr[5], Amy Ciccarella[6], Ana Calcagnotto[5], Peter G. Mikhael[1], Samantha J. Mentch[1], Juan Liu[1], Gene Ables[7], David G. Kirsch[1,2], David S. Hsu[3,4], Sailendra N. Nichenametla[7] & Jason W. Locasale[1]*

Nutrition exerts considerable effects on health, and dietary interventions are commonly used to treat diseases of metabolic aetiology. Although cancer has a substantial metabolic component[1], the principles that define whether nutrition may be used to influence outcomes of cancer are unclear[2]. Nevertheless, it is established that targeting metabolic pathways with pharmacological agents or radiation can sometimes lead to controlled therapeutic outcomes. By contrast, whether specific dietary interventions can influence the metabolic pathways that are targeted in standard cancer therapies is not known. Here we show that dietary restriction of the essential amino acid methionine—the reduction of which has anti-ageing and anti-obesogenic properties—influences cancer outcome, through controlled and reproducible changes to one-carbon metabolism. This pathway metabolizes methionine and is the target of a variety of cancer interventions that involve chemotherapy and radiation. Methionine restriction produced therapeutic responses in two patient-derived xenograft models of chemotherapy-resistant RAS-driven colorectal cancer, and in a mouse model of autochthonous soft-tissue sarcoma driven by a G12D mutation in KRAS and knockout of p53 (*Kras*$^{G12D/+}$;*Trp53*$^{-/-}$) that is resistant to radiation. Metabolomics revealed that the therapeutic mechanisms operate via tumour-cell-autonomous effects on flux through one-carbon metabolism that affects redox and nucleotide metabolism—and thus interact with the antimetabolite or radiation intervention. In a controlled and tolerated feeding study in humans, methionine restriction resulted in effects on systemic metabolism that were similar to those obtained in mice. These findings provide evidence that a targeted dietary manipulation can specifically affect tumour-cell metabolism to mediate broad aspects of cancer outcome.

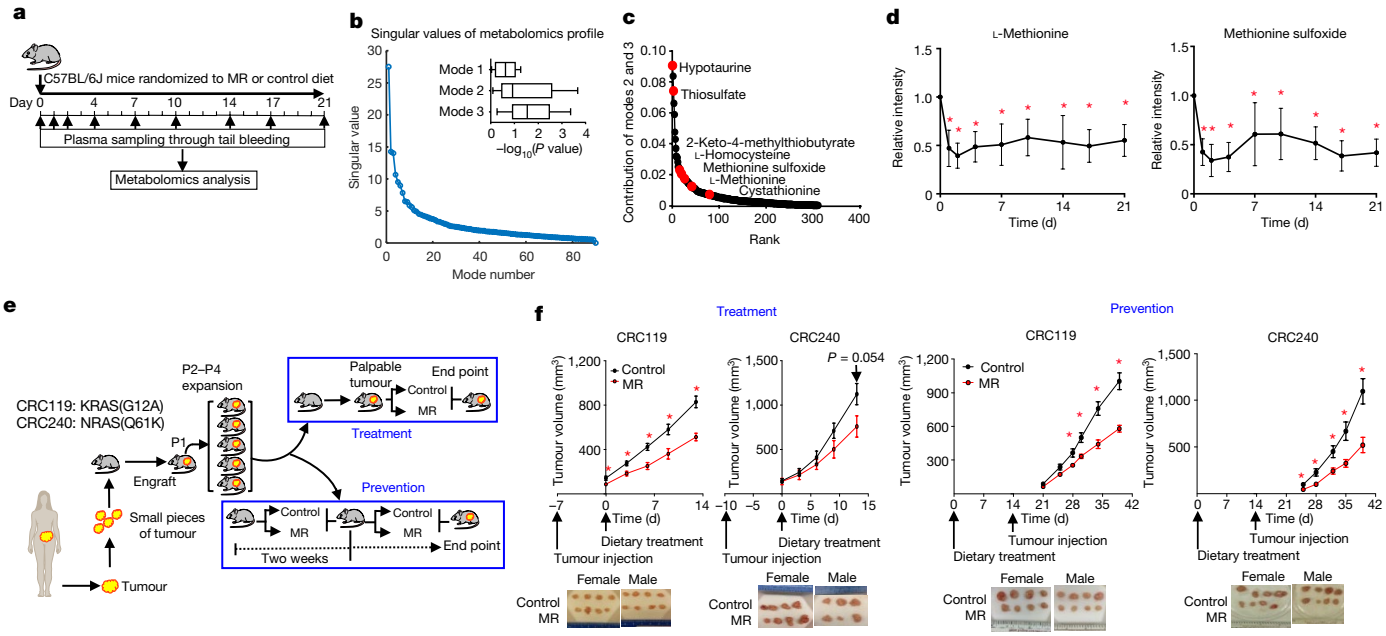


Fig. 1 | Dietary methionine restriction rapidly and specifically alters methionine and sulfur metabolism and inhibits tumour growth in PDX models of colorectal cancer. a, Experimental design in C57BL/6J mice. $n = 5$ mice per group. MR, methionine restriction. **b**, Ninety sets of metabolic profiles from **a** were computed for singular values via singular value decomposition. Insert, two-sided *t*-test *P* values assessing the difference between control and methionine restriction in the first three modes. For definitions of the modes, see 'Analysis of the time-course metabolomics data' in Methods. $n = 5$ mice per group. Box limits are the 25th and 75th percentiles, centre lines are median, and the whiskers are the minimal and maximal values. **c**, Contribution of modes 2 and 3 in **b** ranked across all measured metabolites. **d**, Relative intensity of methionine and methionine sulfoxide. Mean ± s.d., $n = 5$ mice per group, $^*P < 0.05$ by two-tailed Student's *t*-test. **e**, Schematic of experimental design using colorectal PDXs. Treatment, $n = 7$ mice per group (4 female and 3 male). Prevention, $n = 8$ mice per group (4 female and 4 male). P1, passage 1; P2, passage 2; P3, passage 3; P4, passage 4. **f**, Tumour growth curve and images of tumours at the end point from **e**. Mean ± s.e.m., $^*P < 0.05$ by two-tailed Student's *t*-test.

[1]Department of Pharmacology and Cancer Biology, Duke University School of Medicine, Durham, NC, USA. [2]Department of Radiation Oncology, Duke University Medical Center, Durham, NC, USA. [3]Center for Genomics and Computational Biology, Duke University, Durham, NC, USA. [4]Department of Medical Oncology, Duke University Medical Center, Durham, NC, USA. [5]Department of Public Health Sciences, Penn State University College of Medicine, Hershey, PA, USA. [6]Penn State University Clinical Research Center, State College, PA, USA. [7]Orentreich Foundation for the Advancement of Science, Cold Spring, NY, USA. [8]These authors contributed equally: Sydney M. Sanderson, Ziwei Dai. *e-mail: Jason.Locasale@duke.edu

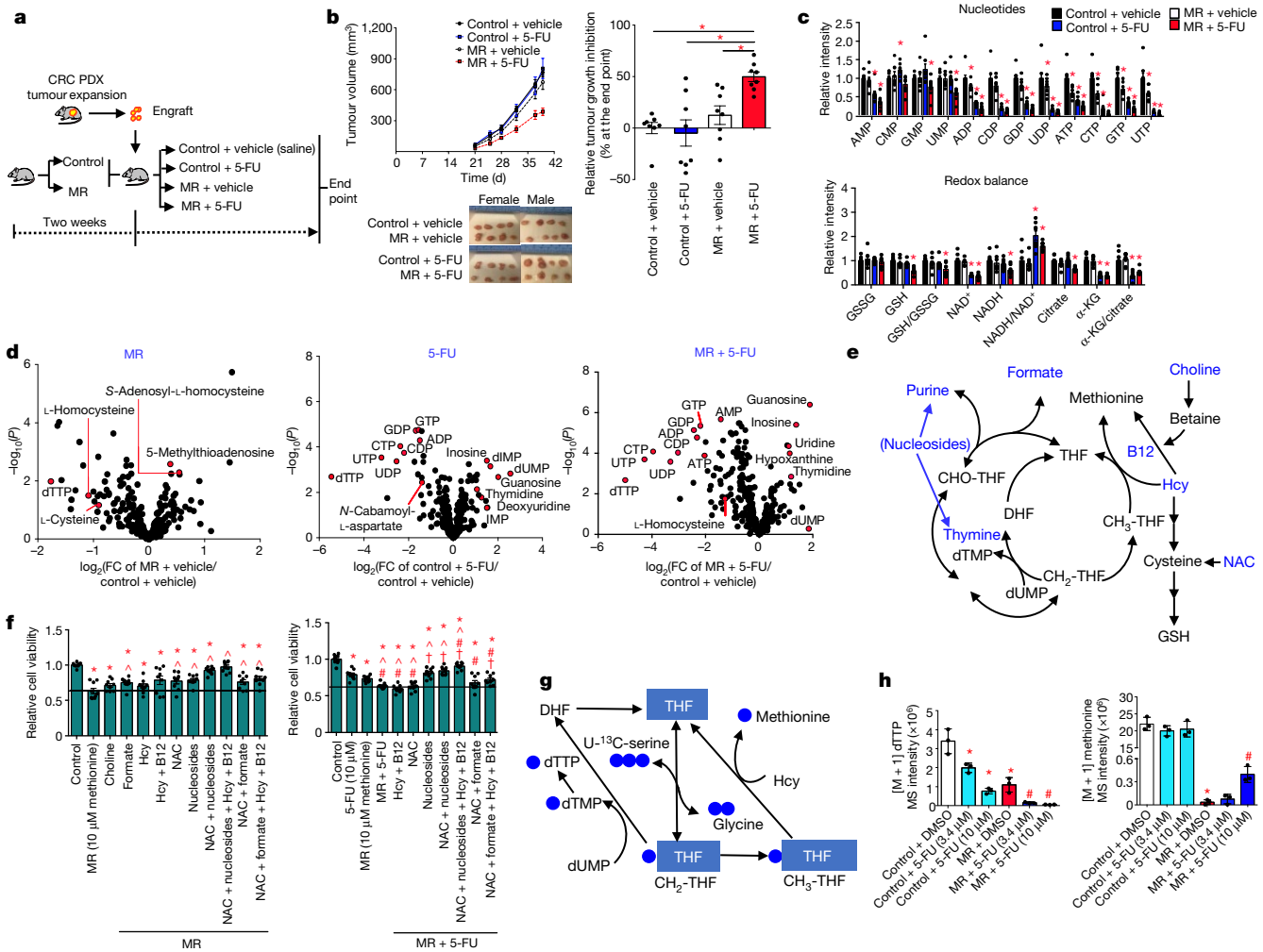


Fig. 2 | Dietary methionine restriction sensitizes PDX models of colorectal cancer to chemotherapy with 5-FU. a, Experimental design. CRC, colorectal cancer. **b**, Tumour growth curves, quantification and images at the end point. Mean ± s.e.m., $*P < 0.05$ by two-tailed Student's *t*-test. $n = 8$ mice per group (4 female and 4 male). **c**, Relative intensity of metabolites related to nucleotide metabolism and redox balance in tumours. Mean ± s.e.m., $*P < 0.05$ versus control by two-tailed Student's *t*-test. $n = 8$ mice per group. α-KG, α-ketoglutarate; GSH, glutathione; GSSG, the oxidized form of glutathione. **d**, Volcano plots of metabolites in tumours. FC, fold change. *P* values were determined by two-tailed Student's *t*-test. **e**, Schematic of supplementation experiments, with added metabolites in blue. B12, vitamin B12; Hcy, homocysteine. **f**, Effect of nutrient supplements on methionine restriction alone or with 5-FU-inhibited cell proliferation in CRC119 primary cells. Mean ± s.e.m., $n = 9$ biologically independent samples from three independent experiments. $*P < 0.05$ versus control, $\wedge P < 0.05$ versus methionine restriction, $\#P < 0.05$ versus 5-FU; $\dagger P < 0.05$ versus methionine restriction + 5-FU by two-tailed Student's *t*-test. **g**, U-^{13}C-serine tracing. **h**, Mass intensity for [M + 1] dTTP and [M + 1] methionine in CRC119 cells. MS, mass spectra. Mean ± s.d., $n = 3$ biologically independent samples. $*P < 0.05$ versus control, $\#P < 0.05$ versus methionine restriction by two-tailed Student's *t*-test.

Nutrient composition in growth media has marked effects on cancer cell metabolism[3–5]. However, the extent to which diet—through its influence on levels of circulating metabolites (which is the in vivo equivalent of medium nutrient composition)—alters metabolic pathways in tumours and affects therapeutic outcomes is largely unknown. Previous studies have shown that the dietary removal of serine and glycine can modulate cancer outcome[6–8]. The availability of histidine and asparagine mediates the response to methotrexate[9] and the progression of breast cancer metastasis[10], respectively. Whether such interventions broadly affect metabolism or have targeted effects on specific pathways related to these nutrients is unknown. One possibility for a specific dietary intervention in cancer is the restriction of methionine, which is an essential amino acid in one-carbon metabolism. Methionine is the most variable metabolite found in human plasma[11], and has a myriad of functions as a result of its location in one-carbon metabolism[12]. Dietary restriction of methionine is known to extend lifespan[13,14] and improve metabolic health[15–17]. One-carbon metabolism, through its essential role in redox and nucleotide metabolism, is the target of frontline cancer chemotherapies such as 5-fluorouracil (5-FU), and radiation therapy[18–20]. Indeed, some cancer cell lines are auxotrophic for methionine[21], and depleting or restricting methionine from the diet may have anti-cancer effects in mice[22–24]. We therefore reasoned that methionine restriction could have broad anti-cancer properties by targeting a focused area of metabolism, and that these anti-cancer effects would interact with the response to other therapies that also affect one-carbon metabolism.

Methionine restriction alters metabolism in mouse liver and plasma after a long-term intervention[11], but its effect on acute time scales has not been explored in as much detail. We switched the diet of C57BL/6J male mice from chow to a control (0.86% methionine, w/w) or a methionine-restricted (0.12% methionine, w/w) diet, and obtained plasma metabolite profiles over time (Fig. 1a). We studied the metabolic dynamics using singular value decomposition (Fig. 1b, Extended Data Fig. 1a) and observed coordinated changes related to methionine and sulfur metabolism (Fig. 1b, c), which were confirmed with hierarchical clustering (Extended Data Fig. 1b). Methionine restriction reduced the levels of methionine-related metabolites within two days, and these levels were sustained throughout the intervention (Fig. 1d, Extended

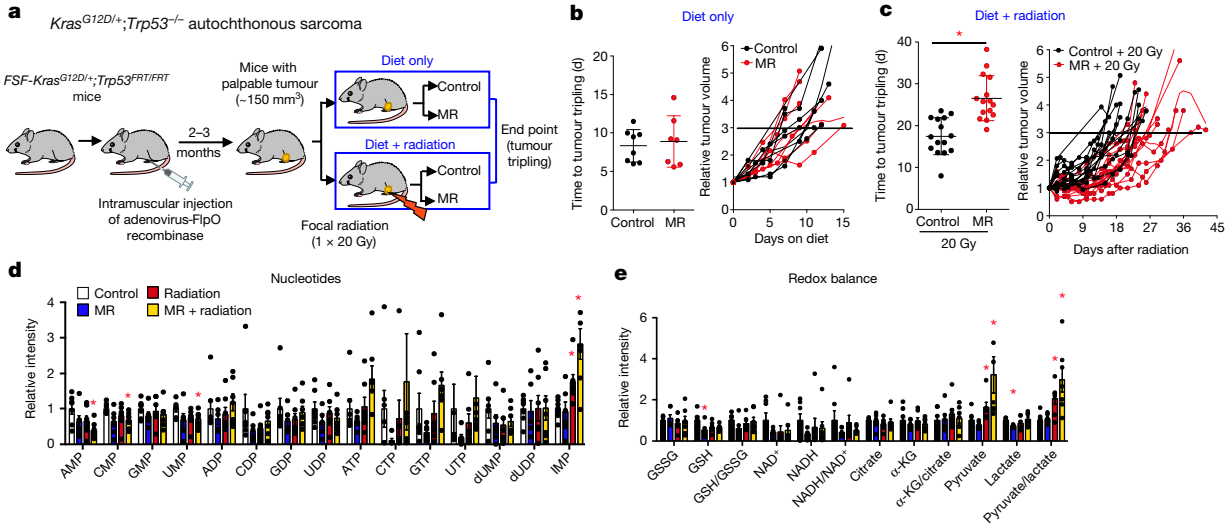


Fig. 3 | Dietary methionine restriction sensitizes mouse models of RAS-driven autochthonous sarcoma to radiation. a, Experimental design. **b**, Time to tumour tripling and tumour growth curve from mice on dietary treatment only. Mean ± s.d., control, $n = 8$ mice; methionine restriction, $n = 7$ mice. **c**, Time to tumour tripling and tumour growth curve from mice on the combination of dietary treatment and radiation. Mean ± s.d., $n = 15$ mice per group. $^*P < 0.05$ by two-tailed Student's t-test. **d**, **e**, Relative intensity of nucleotides (**d**) and metabolites related to redox balance (**e**) in tumours. Mean ± s.e.m., $n = 7$ mice per group, except for methionine restriction ($n = 6$). $^*P < 0.05$ compared to the control group by two-tailed Student's t-test.

Data Fig. 1b–f, Extended Data Table 1). Given these rapid and specific effects, we sought to evaluate methionine restriction in a series of pre-clinical settings that relate to one-carbon metabolism. We first considered two patient-derived xenograft (PDX) models of RAS-driven colorectal cancer, one of which (named CRC119) bears a KRAS(G12A) mutation and the other (named CRC240) a NRAS(Q61K) mutation (Extended Data Fig. 2a). Mice were subjected to the control or methionine-restricted diet when the tumour was palpable (for treatment settings) or two weeks before inoculation (for prevention settings) (Fig. 1e). Methionine restriction inhibited tumour growth in CRC119 ($P = 5.71 \times 10^{-12}$ at the end point, two-tailed Student's t-test), and showed an inhibitory effect in CRC240 ($P = 0.054$ at the end point, two-tailed Student's t-test) (Fig. 1f). Similar or higher amounts of food intake were observed with the methionine-restricted diet relative to the control diet (Extended Data Fig. 2b), which implies the inhibitory effect was not due to caloric restriction. To gain insights into metabolism, we profiled metabolites in tumour, plasma and liver and found that in each case methionine restriction altered methionine and sulfur-related metabolism (Extended Data Fig. 2c–e). A comparative metabolomics analysis across tissues showed that these effects most probably occur in a cell-autonomous manner (Extended Data Fig. 3, Methods), and could be confirmed with data in cell culture (Extended Data Fig. 4). Thus, the inhibition of tumour growth is at least partially (if not largely) attributable to lower circulating levels of methionine, which lead to cell-autonomous effects on tumours.

5-FU targets thymidylate synthase[18] and is a frontline chemotherapy for colorectal cancer, with therapeutic strategies achieving modest (approximately 60–65%) responses[25,26]. We therefore tested whether methionine restriction could synergize with 5-FU in the CRC119 model (Fig. 2a). We delivered a tolerable low dose of 5-FU that alone showed no effect on tumour growth (Fig. 2b). Methionine restriction synergized with 5-FU treatment, leading to a marked inhibition of tumour growth, a broad effect on metabolic pathways in tumour, plasma and liver, and, most prominently, changes to nucleotide metabolism and redox state that were related to both the mechanistic action of 5-FU and methionine restriction (Fig. 2b–d, Extended Data Fig. 5a–g). Fold changes of metabolites were highly correlated between plasma and liver (Spearman's rho = 0.38, $P = 6.7 \times 10^{-11}$) but not between tumour and liver (Spearman's rho = 0.14, $P = 0.02$) or circulation (Spearman's rho = 0.14, $P = 0.03$), which indicates that methionine restriction exerted specific effects on tumours (Extended Data Fig. 5h). Dietary restriction of methionine therefore synergizes with 5-FU, inhibiting the growth of colorectal cancer tumours and disrupting nucleotide metabolism and redox balance.

Next, we supplemented primary CRC119 cells and HCT116 colorectal cancer cells with nutrients related to methionine metabolism, in the presence of methionine restriction, 5-FU or both (Fig. 2e, Extended Data Fig. 6a, b). Nucleosides and *N*-acetylcysteine (NAC), along with related supplements, partially alleviated the inhibition of cell proliferation due to methionine restriction, both with and without 5-FU treatment, in CRC119 cells (Fig. 2f). These observations were largely replicated in HCT116 cells (Extended Data Fig. 6b). Using serine uniformly labelled with ^{13}C (U-^{13}C-serine), we found that methionine restriction and 5-FU led to a further reduction of [M + 1] dTTP caused by 5-FU, with an increase of [M + 1] methionine (Fig. 2g, h, Extended Data Fig. 6c). Thus, the synergistic effect between methionine restriction and 5-FU treatment is at least partially due to an increase in methionine synthesis, which competes with dTMP synthesis for the serine-derived one-carbon unit 5,10-methylene-tetrahydrofolate. These data support the conclusion that disruption to nucleotide metabolism and redox balance contributes to the inhibition of cell proliferation that is induced by methionine restriction.

To further explore the therapeutic potential of dietary restriction of methionine and related mechanisms, we considered an autochthonous mouse model of radiation resistance in soft-tissue sarcoma[27]. Extremity sarcomas were induced in *FSF-Kras*$^{G12D/+}$;*Trp53*$^{FRT/FRT}$ mice within two to three months of intramuscular delivery of adenovirus that expresses FlpO recombinase (Fig. 3a, Methods). The autochthonous and PDX models together span the spectrum of acceptable pre-clinical tumour models, and these cancer types allow for the investigation of treatments related to one-carbon metabolism (that is, chemotherapy in colorectal cancer and radiation in sarcoma). Methionine restriction alone did not alter tumour growth in this aggressive autochthonous model, and led to minimal effects on methionine metabolism (Fig. 3b, Extended Data Fig. 7a, b). Methionine restriction with a focal dose of radiation (20 Gy) reduced tumour growth and extended the tumour tripling time by 52%, from an average of 17.48 days to 26.57 days (Fig. 3c), which is comparable to effects seen with known radiosensitizing agents[28]. These effects appeared to be tumour-cell-autonomous and not attributable to protein synthesis or methylation reactions (Extended Data Fig. 7c–e). Nevertheless, disruptions to nucleotide- and redox-related metabolism

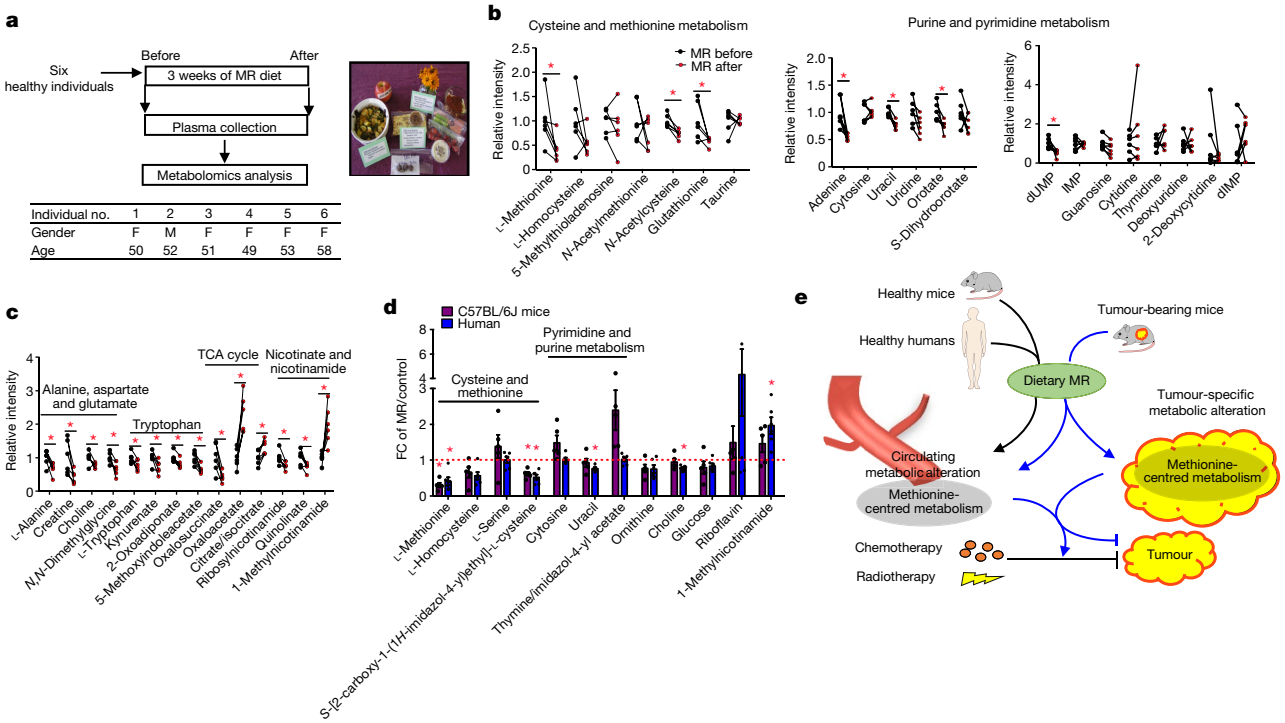


Fig. 4 | Dietary methionine restriction can be achieved in humans. **a**, Experimental design, including background information on participants in the dietary study and representative daily methionine-restricted diet. **b**, **c**, Relative intensity of plasma metabolites related to cysteine and methionine metabolism, and purine and pyrimidine metabolism (**b**), and in the other most affected pathways (**c**). $n = 6$ individuals. $^*P < 0.05$ by two-tailed Student's t-test. **d**, Methionine-restriction-induced fold changes of plasma metabolites in cysteine and methionine metabolism, and pyrimidine and purine metabolism in C57BL/6J mice ($n = 5$) and humans ($n = 6$). Mean ± s.e.m. $^*P < 0.05$ by two-tailed Student's t-test. **e**, Model of the influence of dietary methionine restriction on tumour-cell metabolism.

were observed and may underlie the effects of methionine restriction in combination with radiation (Fig. 3d, e, Extended Data Fig. 7c–g).

Finally, in a proof-of-principle clinical study, we recruited six healthy middle-aged individuals and subjected them to a low methionine diet (about 2.92 mg kg^{-1} day^{-1})—equivalent to an 83% reduction in daily methionine intake—for three weeks (Fig. 4a, Methods). Methionine restriction reproducibly suppressed methionine levels and altered circulating metabolism, with cysteine and methionine metabolism among the top altered metabolic pathways (Fig. 4b, Extended Data Fig. 8a–c). Methionine restriction reduced NAC and glutathione in all subjects, and affected metabolites related to methylation, nucleotide metabolism, the tricarboxylic acid cycle and amino acid metabolism (Fig. 4b, c, Extended Data Fig. 8d). Plasma methionine-related metabolites in healthy humans were highly correlated with those in all mouse models (Spearman's rho = 0.53–0.73) (Extended Data Fig. 9a, b, Fig. 4d), which indicates that the response to methionine restriction is conserved between humans and mice. This controlled clinical study extends observations obtained from studies using methionine-free diets that are toxic[29,30] to methionine restriction at levels that are tolerated in humans, and provide reasonable dietary possibilities—including levels of methionine that may be possible to obtain with vegan or some Mediterranean diets.

Together, we provide evidence that dietary restriction of methionine induces rapid and specific metabolic profiles in mice and humans that can be induced in a clinical setting. By disrupting the flux backbone of one-carbon metabolism with methionine restriction, vulnerabilities involving redox and nucleotide metabolism are created and can be exploited by administration of other therapies (here, radiation and antimetabolite chemotherapy) that target these aspects of cancer metabolism (Fig. 4e). Thus, a synthetic lethal interaction is defined with the diet and the otherwise-resistant treatment modality. This study may help to further establish principles of how dietary interventions may be used to influence cancer outcomes in broader contexts.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1437-3.

Received: 27 July 2018; Accepted: 26 June 2019;
Published online 31 July 2019.

1. DeBerardinis, R. J. & Chandel, N. S. Fundamentals of cancer metabolism. *Sci. Adv.* **2**, e1600200 (2016).
2. Goncalves, M. D., Hopkins, B. D. & Cantley, L. C. Dietary fat and sugar in promoting cancer development and progression. *Annual Rev. Cancer Biol.* **3**, 255–273 (2019).
3. Cantor, J. R. et al. Physiologic medium rewires cellular metabolism and reveals uric acid as an endogenous inhibitor of UMP synthase. *Cell* **169**, 258–272 (2017).
4. Tardito, S. et al. Glutamine synthetase activity fuels nucleotide biosynthesis and supports growth of glutamine-restricted glioblastoma. *Nat. Cell Biol.* **17**, 1556–1568 (2015).
5. Liu, X., Romero, I. L., Litchfield, L. M., Lengyel, E. & Locasale, J. W. Metformin targets central carbon metabolism and reveals mitochondrial requirements in human cancers. *Cell Metab.* **24**, 728–739 (2016).
6. Maddocks, O. D. et al. Serine starvation induces stress and p53-dependent metabolic remodelling in cancer cells. *Nature* **493**, 542–546 (2013).
7. Maddocks, O. D. K. et al. Modulating the therapeutic response of tumours to dietary serine and glycine starvation. *Nature* **544**, 372–376 (2017).
8. Gravel, S. P. et al. Serine deprivation enhances antineoplastic activity of biguanides. *Cancer Res.* **74**, 7521–7533 (2014).
9. Kanarek, N. et al. Histidine catabolism is a major determinant of methotrexate sensitivity. *Nature* **559**, 632–636 (2018).
10. Knott, S. R. V. et al. Asparagine bioavailability governs metastasis in a model of breast cancer. *Nature* **554**, 378–381 (2018).
11. Mentch, S. J. et al. Histone methylation dynamics and gene regulation occur through the sensing of one-carbon metabolism. *Cell Metab.* **22**, 861–873 (2015).
12. Gao, X., Reid, M. A., Kong, M. & Locasale, J. W. Metabolic interactions with cancer epigenetics. *Mol. Aspects Med.* **54**, 50–57 (2017).
13. Orentreich, N., Matias, J. R., DeFelice, A. & Zimmerman, J. A. Low methionine ingestion by rats extends life span. *J. Nutr.* **123**, 269–274 (1993).

14. Lee, B. C. et al. Methionine restriction extends lifespan of *Drosophila melanogaster* under conditions of low amino-acid status. *Nat. Commun.* **5**, 3592 (2014).
15. Malloy, V. L. et al. Methionine restriction prevents the progression of hepatic steatosis in leptin-deficient obese mice. *Metabolism* **62**, 1651–1661 (2013).
16. Ables, G. P., Perrone, C. E., Orentreich, D. & Orentreich, N. Methionine-restricted C57BL/6J mice are resistant to diet-induced obesity and insulin resistance but have low bone density. *PLoS ONE* **7**, e51357 (2012).
17. Malloy, V. L. et al. Methionine restriction decreases visceral fat mass and preserves insulin action in aging male Fischer 344 rats independent of energy restriction. *Aging Cell* **5**, 305–314 (2006).
18. Ser, Z. et al. Targeting one carbon metabolism with an antimetabolite disrupts pyrimidine homeostasis and induces nucleotide overflow. *Cell Reports* **15**, 2367–2376 (2016).
19. Miousse, I. R. et al. One-carbon metabolism and ionizing radiation: a multifaceted interaction. *Biomol. Concepts* **8**, 83–92 (2017).
20. Locasale, J. W. Serine, glycine and one-carbon units: cancer metabolism in full circle. *Nat. Rev. Cancer* **13**, 572–583 (2013).
21. Hoffman, R. M. & Erbe, R. W. High *in vivo* rates of methionine biosynthesis in transformed human and malignant rat cells auxotrophic for methionine. *Proc. Natl Acad. Sci. USA* **73**, 1523–1527 (1976).
22. Komninou, D., Leutzinger, Y., Reddy, B. S. & Richie, J. P. Jr. Methionine restriction inhibits colon carcinogenesis. *Nutr. Cancer* **54**, 202–208 (2006).
23. Hens, J. R. et al. Methionine-restricted diet inhibits growth of MCF10AT1-derived mammary tumors by increasing cell cycle inhibitors in athymic nude mice. *BMC Cancer* **16**, 349 (2016).
24. Guo, H. et al. Therapeutic tumor-specific cell cycle block induced by methionine starvation *in vivo*. *Cancer Res.* **53**, 5676–5679 (1993).
25. Saltz, L. B. et al. Bevacizumab in combination with oxaliplatin-based chemotherapy as first-line therapy in metastatic colorectal cancer: a randomized phase III study. *J. Clin. Oncol.* **26**, 2013–2019 (2008).
26. Douillard, J. Y. et al. Randomized, phase III trial of panitumumab with infusional fluorouracil, leucovorin, and oxaliplatin (FOLFOX4) versus FOLFOX4 alone as first-line treatment in patients with previously untreated metastatic colorectal cancer: the PRIME study. *J. Clin. Oncol.* **28**, 4697–4705 (2010).
27. Kirsch, D. G. et al. A spatially and temporally restricted mouse model of soft tissue sarcoma. *Nat. Med.* **13**, 992–997 (2007).
28. Moding, E. J. et al. Tumor cells, but not endothelial cells, mediate eradication of primary sarcomas by stereotactic body radiation therapy. *Sci. Transl. Med.* **7**, 278ra34 (2015).
29. Durando, X. et al. Optimal methionine-free diet duration for nitrourea treatment: a phase I clinical trial. *Nutr. Cancer* **60**, 23–30 (2008).
30. Durando, X. et al. Dietary methionine restriction with FOLFOX regimen as first line therapy of metastatic colorectal cancer: a feasibility study. *Oncology* **78**, 205–209 (2010).

METHODS

No statistical methods were used to predetermine sample size.

Animals, diets and tissue collection. All animal procedures and studies were approved by the Institutional Animal Care and Use Committee (IACUC) at Duke University. All experiments were performed in accordance with relevant guidelines and regulations. All mice were housed at 20 ± 2 °C with 50 ± 10% relative humidity and a standard 12-h dark–12-h light cycle. The special diets with defined methionine levels that have previously been used[11,16] were purchased from Research Diets; the control diet contained 0.86% methionine (w/w, catalogue no. A11051302) and methionine-restricted diet contained 0.12% methionine (w/w, catalogue no. A11051301). Three mouse models were used, and are described in 'PDX models of colorectal cancer' and 'Autochthonous soft-tissue sarcomas'. For all animal studies, mice were randomized to the control or methionine-restricted diet, and investigators were not blinded to allocation during experiments or outcome assessment.

Methionine-restriction time-course study in healthy mice. Twelve-week-old male C57BL/6J mice (Jackson Laboratories) were subjected to either the control or the methionine-restricted diet ad libitum for three weeks. Mouse blood was sampled through tail bleeding in the morning (10:00–12:00) at days 1, 2, 4, 7, 10, 14, 17 and 21 after the dietary treatments. By day 21, all mice were euthanized for tissue collection.

PDX models of colorectal cancer. PDX models of colorectal cancer with liver metastasis were developed as previously described[31,32], under an IRB-approved protocol (Pro00002435). In brief, CRC119 and CRC240 tumours were resected, washed and minced, and then passaged through JAX *NOD.CB17-PrkdcSCID-J* mice 2–5 times. For the dietary studies, CRC119 and CRC240 PDX tumours were minced in PBS at 150 mg/ml and 200 μl of tumour suspension was subcutaneously injected into the flanks of NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ mice from the Jackson Laboratory. Mice (four female and three or four male) were subjected to the control or methionine-restricted diet, either two weeks before the tumour injection or from when the tumour was palpable until the end point (a tumour volume of about 1,500 mm^3). Tumour size was monitored two to three times per week until the end point. For the combination therapy with the standard chemotherapy drug 5-FU, mice were subjected to the control or the methionine-restricted diet from two weeks before the tumour injection until the end point. When tumours were palpable, mice (four female and four male) were randomized to treatment of 5-FU (NDC 63323-117-10, 12.5 mg/kg three times per week) or vehicle (saline) through intraperitoneal injection. To minimize toxicity, we delivered an established low dose of 5-FU[33]. Tumour size was monitored two to three times per week until the end point.

Autochthonous soft-tissue sarcomas. Primary soft-tissue sarcomas were generated as previously described[27,34]. In brief, *Trp53*^{FRT} mice were crossed with mice that carry an Flp-activated allele of oncogenic *Kras* (*FSF-Kras*^{G12D}) to generate *FSF-Kras*$^{G12D/+}$;*Trp53*$^{FRT/FRT}$ compound conditional-mutant mice (KP mice). *Trp53*^{FRT} mice and *FSF-Kras*^{G12D} mice were maintained on mixed C57BL/6J × 129SvJ backgrounds. Soft-tissue sarcomas were induced by intramuscular injection of an adenovirus that expresses FlpO into KP mice. Twenty-five microlitres of Ad5CMVFlpO (6×10^{10} plaque-forming units per millilitre) was incubated with 600 μl minimum essential medium (Sigma-Aldrich) and 3 μl 2 M $CaCl_2$ (Sigma-Aldrich) for 15 min to form calcium phosphate precipitates. Fifty microlitres of precipitated virus was injected intramuscularly per mouse to generate sarcomas. Soft-tissue sarcomas developed at the site of injection in the lower extremity as early as two months after injection. *FSF-Kras*^{G12D} mice were provided by T. Jacks at MIT, and *Trp53*^{FRT} mice had previously been generated at Duke University[27,34].

KP mice (of mixed sex) were subjected to a control or methionine-restricted diet when tumours were palpable (about 150 mm^3), until the end point when the tumour tripled in size. Tumour size was monitored two or three times per week. For combination therapy with radiation, KP mice with palpable tumours were subjected to a single dose of 20 Gy focal radiation, which is moderately effective in this model[35], using the X-RAD 225Cx small-animal image-guided irradiator (Precision X-Ray). The irradiation field was centred on the target via fluoroscopy with 40-kilovolt peak (kVp), 2.5-mA X-rays using a 2-mm aluminium filter. Sarcomas were irradiated with parallel-opposed anterior and posterior fields with an average dose rate of 300 cGy/min prescribed to midplane with 225-kVp, 13-mA X-rays using a 0.3-mm copper filter and a collimator with a 40 × 40-mm^2 radiation field at the treatment isocentre. The dose rate was monitored in an ion chamber by members of the Radiation Safety Division at Duke University. After radiation, mice were immediately subjected to the control or the methionine-restricted diet until the end point at which the tumour tripled in size. Tissues (tumour, liver and plasma) were collected at the time of tumour tripling. For metabolomics analysis, another cohort of mice on the combination therapy with radiation was euthanized at ten days after the radiation and dietary treatment (the average time point at which the tumour size tripled in the KP mice on the control or the methionine-restricted diet alone).

Tissue collection. For tissue collection from all the above mouse studies, mice were fasted in the morning for four hours (9:00–13:00). Tumour, plasma and liver were collected and then immediately snap-frozen, and stored at −80 °C until processed.

Colorectal cancer cell lines. Early-passage primary CRC119 and CRC240 colorectal cancer cell lines were developed from the PDXs. PDXs were collected and homogenized, and the homogenates were grown in RPMI 1640 medium with addition of 10% fetal calf serum, 100,000 U/l penicillin and 100 mg/l streptomycin at 5% CO_2. A single-cell clone was isolated using an O ring. The HCT116 cell line was a gift from the laboratory of L. Cantley, and was maintained in RPMI 1640 supplemented with 10% fetal bovine serum and 100,000 U/l penicillin and 100 mg/l streptomycin. Cells were grown at 37 °C with 5% CO_2. Cell lines were authenticated and tested for mycoplasma at the Duke University DNA Analysis Facility by analysing DNA samples from each cell lines for polymorphic short tandem repeat markers using the GenePrint 10 kit from Promega. All cell lines were negative for mycoplasma contamination.

Cell viability assay. Cell viability was determined by MTT (Invitrogen) assays. In brief, cells cultured in 96-well plates were incubated in RPMI medium containing MTT (final concentration 0.5 mg/ml) in a cell incubator for 2–4 h. The medium was then removed and replaced with 100 μl DMSO, followed by additional 10 min of incubation at 37 °C. The absorbance at 540 nm was read using a plate reader. For the metabolite rescue studies, 10 μM methionine (one tenth of the amount in the RPMI medium) was used to approximately model dietary methionine restriction. We evaluated the effects of supplementation of a suite of nutrients related to methionine metabolism, and the observed differences in metabolite profiles of the mouse models of colorectal cancer—including the one-carbon donors choline and formate; the sulfur donor homocysteine, with or without cofactor vitamin B12; nucleosides; and the antioxidant NAC—on methionine restriction alone or in combination with 5-FU treatment caused defects in cell proliferation. The following final conditions of metabolites were used: homocysteine (400 μM), vitamin B12 (20 μM), nucleosides (Millipore, 1×), choline (1 mM), formate (0.5 mM), NAC (1 mM) and 5-FU (2.5 μM).

Human dietary study. The controlled feeding study was conducted at Penn State University Clinical Research Center and approved by the IRB of the Penn State College of Medicine, in accordance with the Helsinki Declaration of 1975 as revised in 1983 (IRB no. 32378). We have complied with all relevant ethical regulations. Healthy adults of mixed gender were recruited by fliers and word of mouth and—as assessed for initial eligibility by telephone interview—were free of disease and not currently taking specific medications (including anti-inflammatory drugs, corticosteroids, statins, thyroid drugs and oral contraceptives). Final eligibility was assessed by standard clinical chemistry and haematology analyses. Written consent was obtained from eligible subjects and baseline resting metabolic rate was assessed by indirect calorimetry (Parvo Medics), physical activity by questionnaire and dietary intake by 3 unannounced 24-h diet recalls conducted by telephone in the week before returning to the clinic. All of the subjects—6 healthy adults (5 women and 1 man) with a mean ± s.d. age of 52.2 ± 3.19 years (range of 49–58 years) and body mass index of 27.6 ± 4.32 kg/m^2—were placed on an methionine-restricted diet for the final 3 weeks, which provided 50–53% of energy from carbohydrate, 35–38% from fat and 12–13% from protein; total calories were adjusted individually on the basis of the baseline of resting metabolic rate and physical activity (calculated by the Harris Benedict Equation). Of the total protein, 75% was provided by a methionine-free medicinal beverage (Hominex-2, Abbott Nutrition) and the remaining 25% was from low-methionine foods such as fruits, vegetables and refined grains. The total methionine intake was about 2.92 mg kg^{-1} day^{-1}, which represented an 83% reduction in methionine intake compared to pre-test values from diet recalls (about 17.2 mg kg^{-1} day^{-1} at baseline level). The five-day-cycle menu was created and evaluated for nutrient content using the Nutrition Data System for Research. Blood was sampled by a registered nurse into EDTA tubes in the morning after overnight fasting, at the beginning and end of the diet period. Plasma was obtained after centrifugation at 5,000 r.p.m. for 10 min at 4 °C. All six subjects agreed to have their samples and data used for future research. Biosamples were anonymized by re-coding. There was no registration or pre-registration of this study.

Metabolite profiling and isotope tracing. PDX primary cell lines were seeded in 6-well plates at a density of 2.0×10^5 cells per well. For overall polar metabolite profile, after overnight incubation cells were washed once with PBS and cultured for an additional 24 h with 2 ml of conditional RPMI medium containing 0 μM or 100 μM methionine plus 10% FBS. Cellular metabolites were extracted after incubation. For U-^{13}C-serine isotope tracing, both primary CRC119 cells and HCT116 cells were seeded in 6-well plates at a density of 2.0×10^5 cells per well. Cells were washed once with PBS after overnight incubation, and cultured for an additional 24 h with 2 ml of conditional RPMI medium containing 0 μM or 100 μM methionine with or without addition of 5-FU (3.4 or 10 μM) plus 10% FBS. Then, medium was replaced with fresh conditional RPMI medium (0 μM or 100 μM methionine) with or without addition of 5-FU (3.4 or 10 μM) containing tracer U-^{13}C-serine plus

10% dialysed FBS. Cells were traced for 6 h, and tracing was followed by cellular metabolite extraction.

Metabolite extraction. Polar metabolite extraction has previously been described[36]. In brief, tissue samples (liver and tumour) were pulverized in liquid nitrogen and then 3–10 mg of each was weighed out for metabolite extraction using ice-cold extraction solvent (80% methanol/water, 500 μl). Tissue was then homogenized with a homogenizer to an even suspension, and incubated on ice for an additional 10 min. The extract was centrifuged at 20,000*g* for 10 min at 4 °C. The supernatant was transferred to a new Eppendorf tube and dried in vacuum concentrator. For serum or medium, 20 μl of sample was added to 80 μl ice-cold water in an Eppendorf tube on ice, followed by the addition of 400 μl ice-cold methanol. Samples were vortexed at the highest speed for 1 min before centrifugation at 20,000*g* for 10 min at 4 °C. For cells cultured in 6-well plates, cells were placed on top of dry ice right after medium removal. One millilitre ice-cold extraction solvent (80% methanol/water) was added to each well and the extraction plate was quenched at −80 °C for 10 min. Cells were then scraped off the plate into an Eppendorf tube. Samples were vortexed and centrifuged at 20,000*g* for 10 min at 4 °C. The supernatant was transferred to a new Eppendorf tube and dried in vacuum concentrator. The dry pellets were stored at −80 °C for liquid chromatography with high-resolution mass spectrometry analysis. Samples were reconstituted into 30–60 μl sample solvent (water:methanol:acetonitrile, 2:1:1, v/v/v) and were centrifuged at 20,000*g* at 4 °C for 3 min. The supernatant was transferred to liquid chromatography vials. The injection volume was 3 μl for hydrophilic interaction liquid chromatography (HILIC), which is equivalent to a metabolite extract of 160 μg tissue injected on the column.

High-performance liquid chromatography. An Ultimate 3000 UHPLC (Dionex) was coupled to Q Exactive-Mass spectrometer (QE-MS, Thermo Scientific) for metabolite separation and detection. For additional polar metabolite analysis, a HILIC method was used, with an Xbridge amide column (100 × 2.1 mm internal diameter, 3.5 μm; Waters), for compound separation at room temperature. The mobile phase and gradient information has previously been described[37].

Mass spectrometry and data analysis. The QE-MS was equipped with a HESI probe, and the relevant parameters were: heater temperature, 120 °C; sheath gas, 30; auxiliary gas, 10; sweep gas, 3; spray voltage, 3.6 kV for positive mode and 2.5 kV for negative mode. Capillary temperature was set at 320 °C, and the S-lens was 55. A full scan range was set at 60 to 900 (m/z) when coupled with the HILIC method, or 300 to 1,000 (m/z) when low-abundance metabolites needed to be measured. The resolution was set at 70,000 (at m/z 200). The maximum injection time was 200 ms. Automated gain control was targeted at 3 Å ~ 10^6 ions. Liquid chromatography–mass spectrometry peak extraction and integration were analysed with commercially available software Sieve 2.0 (Thermo Scientific). The integrated peak intensity was used for further data analysis. For tracing studies using U-^{13}C-serine, ^{13}C natural abundance was corrected as previously described[38].

Statistical analysis and bioinformatics. Pathway analysis of metabolites was carried out with software Metaboanalyst (http://www.metaboanalyst.ca/MetaboAnalyst/) using the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database (http://www.genome.jp/kegg/). All data are represented as mean ± s.d. or mean ± s.e.m. as indicated. *P* values were calculated by a two-tailed Student's *t* test unless otherwise noted.

Analysis of the time-course metabolomics data. We first constructed a combinational matrix that contained the raw ion intensities of plasma metabolites from C57BL/6J mice (both the control and the methionine-restriction groups). For each group, there were 9 time points and 5 replicates for each time point, which resulted in a 311 × 90 matrix. This matrix was then log-transformed and iteratively row-normalized and column-normalized until the mean values of all rows and columns converged to zero. Singular value decomposition[39] was applied on the processed matrix to identify dominating dynamic modes:

$$A = U\Sigma V^{\mathrm{T}} = \sum_{i=1}^{r} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^{\mathrm{T}}$$

in which A is the processed metabolomics matrix, σ_i is the ith singular value (ranked from maximal to minimal), and $\sigma_i \boldsymbol{v}_i^{\mathrm{T}}$ is termed the ith mode. Elements of $\boldsymbol{u}_i$ (that is, the ith column vector of U) are coefficients for the ith mode. Modes 2 and 3 were defined as responding modes, owing to the significant difference between control and methionine-restriction values in both modes. For the ith metabolite, the total contribution of modes 2 and 3 to its dynamics was evaluated by: $C_{23,i} = \boldsymbol{u}_{i2}^2 + \boldsymbol{u}_{i3}^2$. Mode 1 reflected an overall metabolic change due to switching diets at time zero. Modes 2 and 3 predominantly contained metabolites related to methionine and sulfur metabolism. Time-course metabolomics data of 50 metabolites with highest contribution of modes 2 and 3 were then clustered using the clustergram() function in MATLAB R2018b. All methods used were implemented in MATLAB code. Hierarchical clustering confirmed that a set of methionine-related metabolites was most rapidly suppressed, with other compensatory pathways changing at later times.

Cross-tissue comparison of metabolite fold changes in PDX and sarcoma models. Spearman's rank correlation coefficients were computed on metabolites measured in plasma, tumour and liver. The distance between fold changes in tissues A and B (for example, liver and tumour) was computed by measuring the Euclidean distance between the two vectors of the fold changes that contained all metabolites measured in both A and B. Multidimensional scaling was then applied to visualize the tissues in two dimensions; a stress function, which measures the difference between the dimension-reduced values and the values in the original dataset, was minimized:

$$\min \mathrm{Stress}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N) = \left(\frac{\sum_{ij} (d_{ij} - \|\boldsymbol{x}_i - \boldsymbol{x}_j\|)^2}{\sum_{ij} \boldsymbol{d}_{ij}^2} \right)^{1/2}$$

$$\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N \in \mathbb{R}^2$$

in which N is the total number of metabolites used in the original dataset, $d_{i,j}$ is the Euclidean distance between the ith and jth data points in the original dataset, and $\boldsymbol{x}_i$ is the ith point in the dimension-reduced dataset. All methods used here were implemented in MATLAB.

Methionine-related and methionine-unrelated metabolites. To determine whether the effect of methionine restriction on tumour growth is systemic, cell autonomous or both, we conducted an integrated analysis of global changes in the metabolic network across tumour, plasma and liver within each model from the prevention study in PDX models of colorectal cancer in Fig. 1f. Methionine-related and -unrelated metabolites were defined according to their distance to methionine in the genome-scale metabolic human model Recon 2 (ref. [40]). Metabolites were defined as methionine-related if the distance to methionine was less than or equal to four, or methionine-unrelated when the distance to methionine was larger than four. Metabolites were mapped by their KEGG identity between the metabolomics dataset and Recon 2.

Quantification of methionine concentrations. To quantify methionine concentrations in plasma, liver and tumour across the mouse models and in healthy humans, two additional datasets of metabolomics profiles in human plasma with their corresponding absolute methionine concentrations (quantified using ^{13}C-labelled standards) were used. The raw intensities across all samples were log-transformed and normalized. Linear regression was then performed on the normalized datasets to predict absolute methionine concentrations. Four normalization algorithms including cyclic loess, quantile, median and z-score were tested. Among the normalization algorithms, cyclic loess had the highest R^2 statistics in the corresponding linear regression model ($R^2 = 0.74$ for cyclic loess compared to 0.66 for quantile, 0.68 for median and 0.70 for z-score). Thus, the cyclic-loess-normalized dataset was used for the final model training, which generated the following equation describing the model: $\log(\text{methionine concentration}) = 1.001676 \log(I_{\text{methionine}}) - 14.446017$. In this equation, methionine concentration refers to the absolute methionine concentration, and $I_{\text{methionine}}$ is the cyclic-loess-normalized value of methionine intensity.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The metabolomics data reported in this study have been deposited in Mendeley Data (https://doi.org/10.17632/zs269d9fvb.1).

Code availability

All computer code is available at: https://github.com/LocasaleLab/Dietary_methionine_restriction.

31. Kim, M. K. et al. Characterization of an oxaliplatin sensitivity predictor in a preclinical murine model of colorectal cancer. *Mol. Cancer Ther.* **11**, 1500–1509 (2012).
32. Uronis, J. M. et al. Histological and molecular evaluation of patient-derived colorectal cancer explants. *PLoS ONE* **7**, e38422 (2012).
33. Udofot, O. et al. Pharmacokinetic, biodistribution and therapeutic efficacy of 5-fluorouracil-loaded pH-sensitive PEGylated liposomal nanoparticles in HCT-116 tumor bearing mouse. *J. Nat. Sci.* **2**, e171 (2016).
34. Lee, C. L. et al. Generation of primary tumors with Flp recombinase in *FRT*-flanked *p53* mice. *Dis. Model. Mech.* **5**, 397–402 (2012).
35. Moding, E. J. et al. *Atm* deletion with dual recombinase technology preferentially radiosensitizes tumor endothelium. *J. Clin. Invest.* **124**, 3325–3338 (2014).
36. Liu, X. et al. High-resolution metabolomics with acyl-CoA profiling reveals widespread remodeling in response to diet. *Mol. Cell. Proteomics* **14**, 1489–1500 (2015).

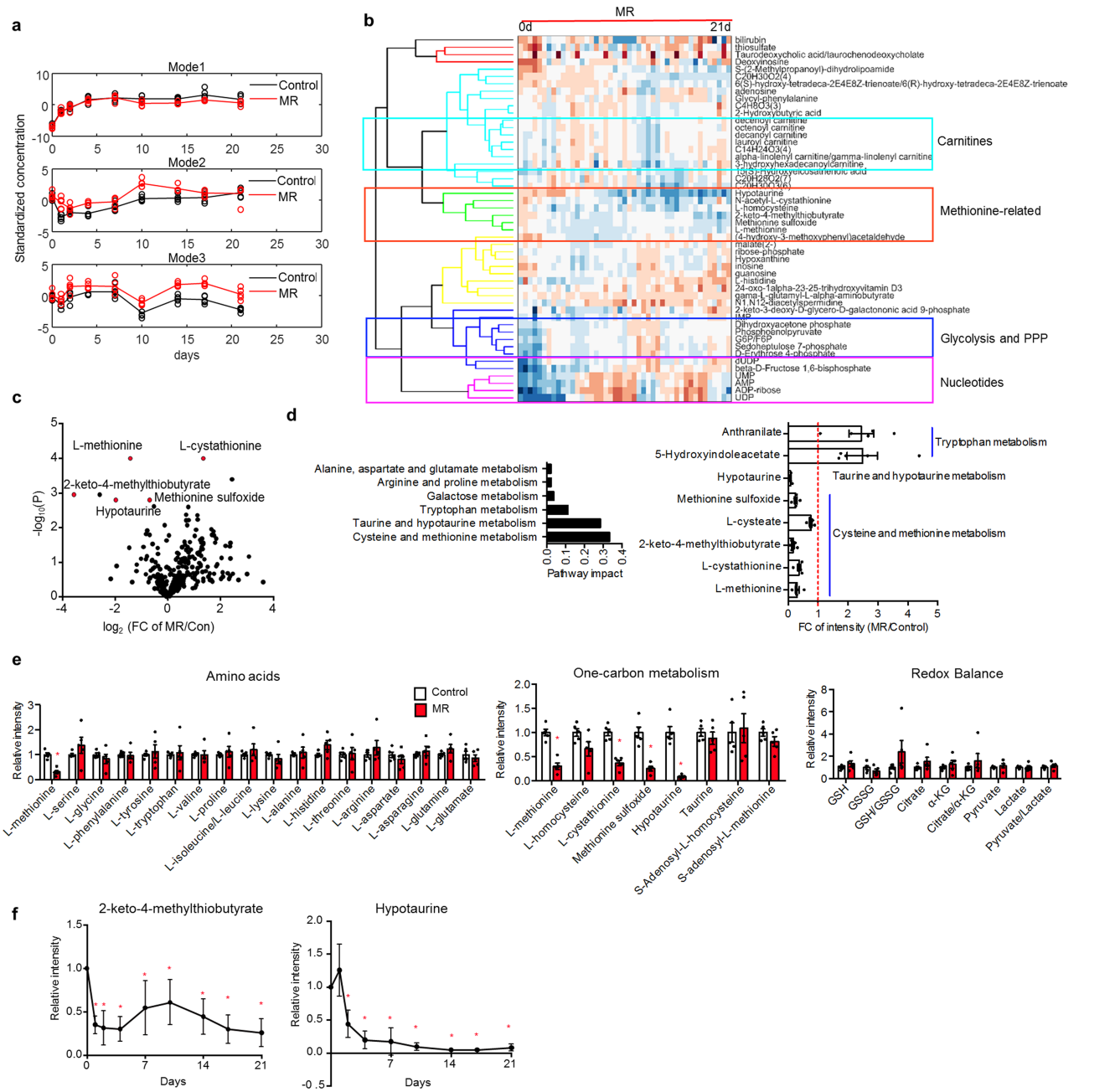
37. Liu, X., Ser, Z. & Locasale, J. W. Development and quantitative evaluation of a high-resolution metabolomics technology. *Anal. Chem*. **86**, 2175–2184 (2014).
38. Yuan, J., Bennett, B. D. & Rabinowitz, J. D. Kinetic flux profiling for quantitation of cellular metabolic fluxes. *Nat. Protocols* **3**, 1328–1340 (2008).
39. Holter, N. S. et al. Fundamental patterns underlying gene expression profiles: simplicity from complexity. *Proc. Natl Acad. Sci. USA* **97**, 8409–8414 (2000).
40. Thiele, I. et al. A community-driven global reconstruction of human metabolism. *Nat. Biotechnol*. **31**, 419–425 (2013).

Acknowledgements We gratefully acknowledge support from the National Institutes of Health (NIH) R01CA193256, R21CA201963 and P30CA014236 (J.W.L.), R35CA197616 (D.G.K.), T32CA93240 (D.E.C.) and the Canadian Institutes of Health Research (CIHR, 146818) (X.G.). We thank M. L. Kiel and T. Hartman for assistance in designing the diets, and S. Heim for help with food preparation in the human study. The human study was partially supported by the Clinical Research Center at Penn State University (NIH M01RR10732). We gratefully acknowledge members of the Locasale laboratory for discussions and apologize to those whose work we could not cite owing to space constraints.

Author contributions X.G. and J.W.L. designed the study, wrote and edited the paper. D.E.C. and D.G.K. designed the sarcoma experiments and edited the paper. M.L. and D.S.H. designed and implemented the colorectal PDX models and edited the paper. X.G., M.L., D.E.C., G.A. and M.A.R. performed animal experiments. X.G., S.M.S. and M.A.R. performed all cell culture experiments. J.P.R. Jr, A. Ciccarella, A. Calcagnotto and S.N.N. conducted the human study. Z.D. conducted computational analyses with initial help from P.G.M. J.L. and S.J.M. assisted in mass spectrometry metabolomics experiments.

Competing interests J.W.L. and X.G. have patents related to targeting amino acid metabolism in cancer therapy. D.G.K. is a co-founder and has equity in XRAD Therapeutics, a company developing radiosensitizing agents. He also has patents related to radiosensitizing agents.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1437-3.
Correspondence and requests for materials should be addressed to J.W.L.
Peer review information *Nature* thanks Danica Chen, Alexei Vazquez and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

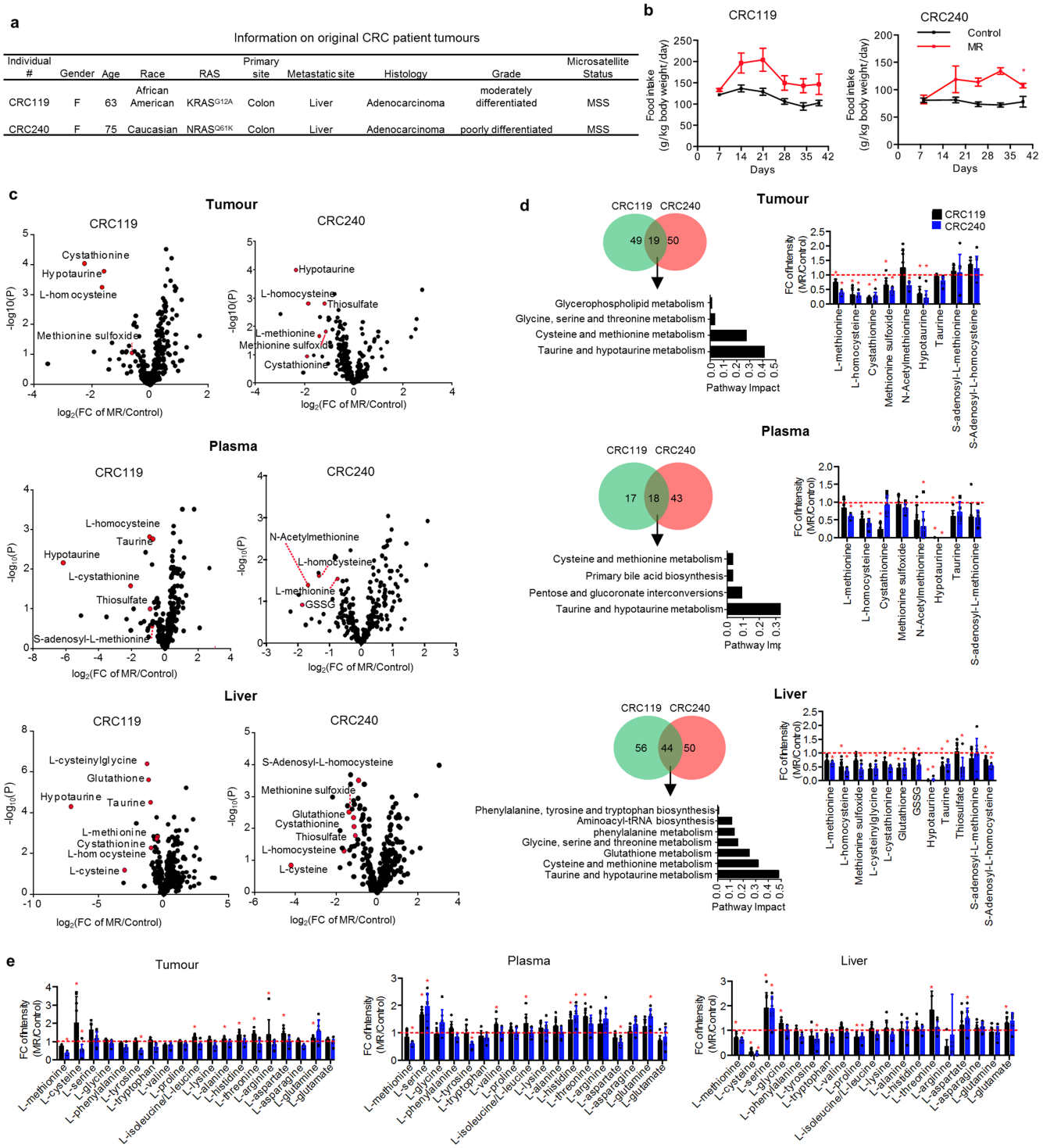


Extended Data Fig. 1 | Dietary restriction of methionine rapidly and specifically alters methionine and sulfur metabolism but maintains overall metabolism in healthy C57BL/6J mice. a, Dynamic patterns of the top three modes. Standardized concentration (the values are normalized to have mean = 0, s.d. = 1) in mode 1, mode 2 and mode 3. **b**, Heat map of metabolites in mode 2 and mode 3. PPP, pentose phosphate pathway. **c**, Volcano plot of metabolites in plasma collected at the end point. *P* values were determined by two-tailed Student's *t*-test. **d**, Left, pathway analysis of significantly changed (*$P < 0.05$, two-tailed Student's *t*-test) plasma metabolites by 21-day methionine-restricted diet. Right, fold change of altered metabolites in the top three most-affected pathways. Mean ± s.e.m., $n = 5$ mice per group. *$P < 0.05$, two-tailed Student's *t*-test. **e**, Relative intensity of plasma amino acids and metabolites in one-carbon metabolism and redox balance at the end of the study. Mean ± s.e.m., $n = 5$ mice per group. *$P < 0.05$, two-tailed Student's *t*-test. **f**, Relative intensity of the methionine-metabolism-related metabolites 2-keto-4-methylthiobutyrate and hypotaurine. Mean ± s.d., $n = 5$ mice per group. *$P < 0.05$, two-tailed Student's *t*-test.

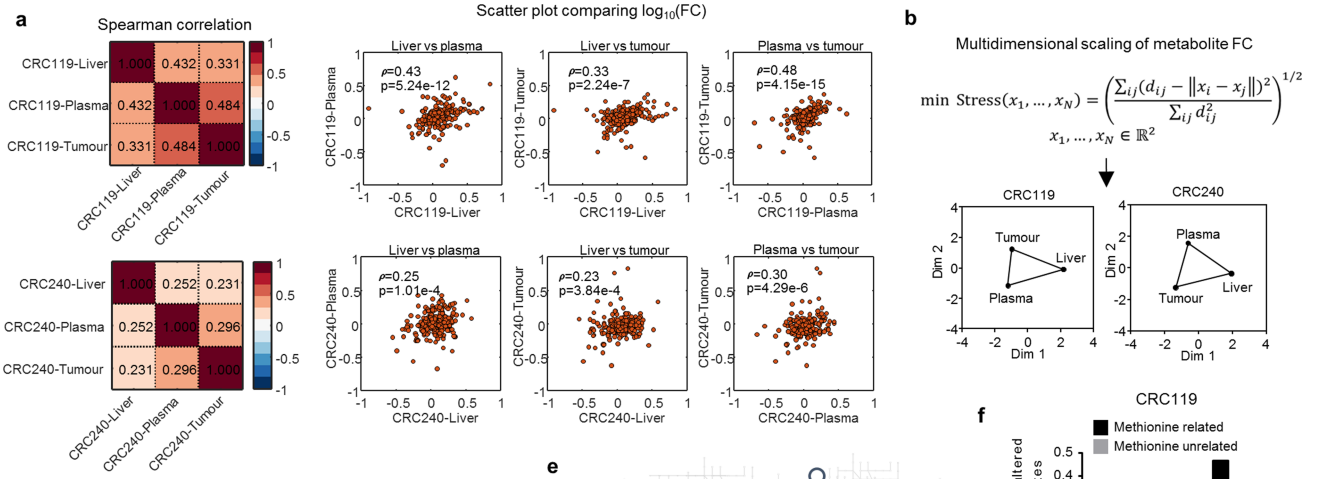
a

Information on original CRC patient tumours

Individual #	Gender	Age	Race	RAS	Primary site	Metastatic site	Histology	Grade	Microsatellite Status
CRC119	F	63	African American	$KRAS^{G12A}$	Colon	Liver	Adenocarcinoma	moderately differentiated	MSS
CRC240	F	75	Caucasian	$NRAS^{Q61K}$	Colon	Liver	Adenocarcinoma	poorly differentiated	MSS



Extended Data Fig. 2 | Dietary restriction of methionine alters methionine metabolism in PDX models of colorectal cancer. **a**, Information on original tumours from patients with colorectal cancer. **b**–**e**, Data from the prevention study shown in Fig. 1f. $n = 8$ mice per group (4 female and 4 male). **b**, Food intake. Mean ± s.e.m. $^*P < 0.05$, two-tailed Student's *t*-test. **c**, Volcano plots of metabolites in tumour, plasma and liver. *P* values were determined by two-tailed Student's *t*-test. **d**, Left, Venn diagrams of significantly changed ($^*P < 0.05$, two-tailed Student's *t*-test) metabolites in tumour, plasma and liver by methionine restriction and pathway analysis (false discovery rate < 0.5) of the commonly changed metabolites. Right, fold changes of intensity of tumour metabolites in cysteine and methionine metabolism, and taurine and hypotaurine metabolism, induced by methionine restriction. Mean ± s.d. $^*P < 0.05$, two-tailed Student's *t*-test. $n = 8$ mice per group (4 female and 4 male). **e**, Relative fold change of intensity of amino acids. Mean ± s.d. $^*P < 0.05$, two-tailed Student's *t*-test. $n = 8$ mice per group (4 female and 4 male).



d

Number of altered metabolites ($p < 0.05$)

Tissue	CRC119	CRC240
Plasma	46	44
Tumour	70	54
Liver	88	76

g

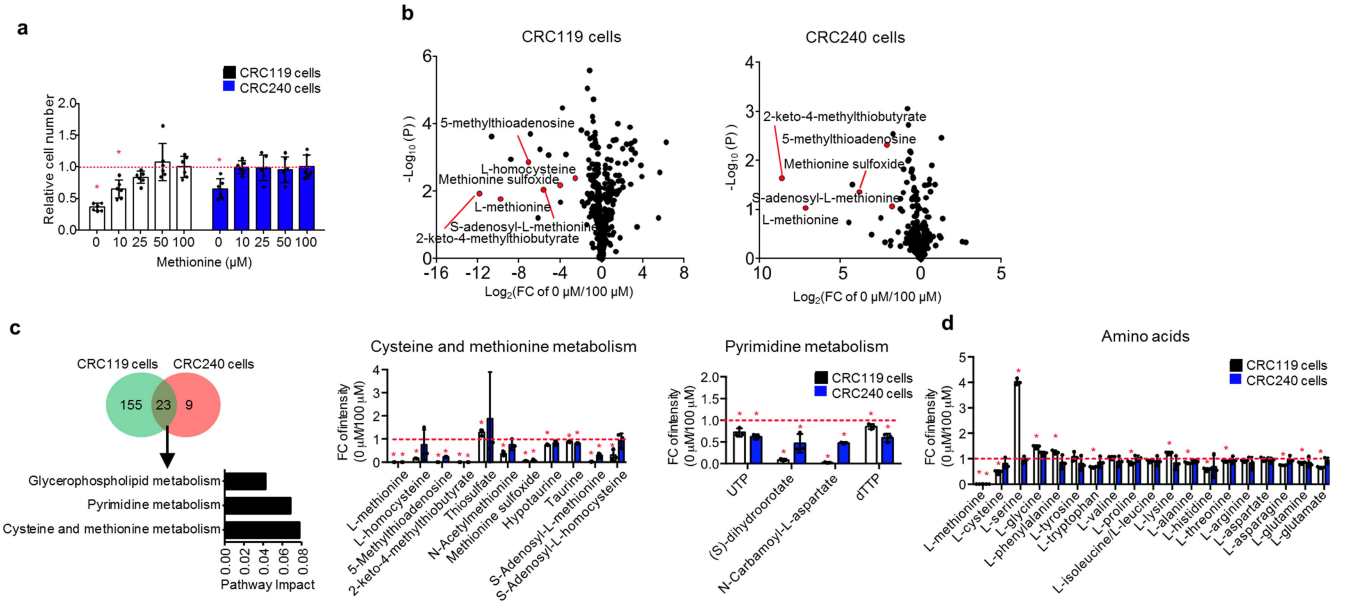
CRC119

Tissue	Methionine related, altered/total	Methionine unrelated, altered/total	p-value by One-sided Fisher's exact test
Plasma	4/17	34/216	0.29
Tumour	7/17	46/216	0.06
Liver	4/17	52/216	0.62

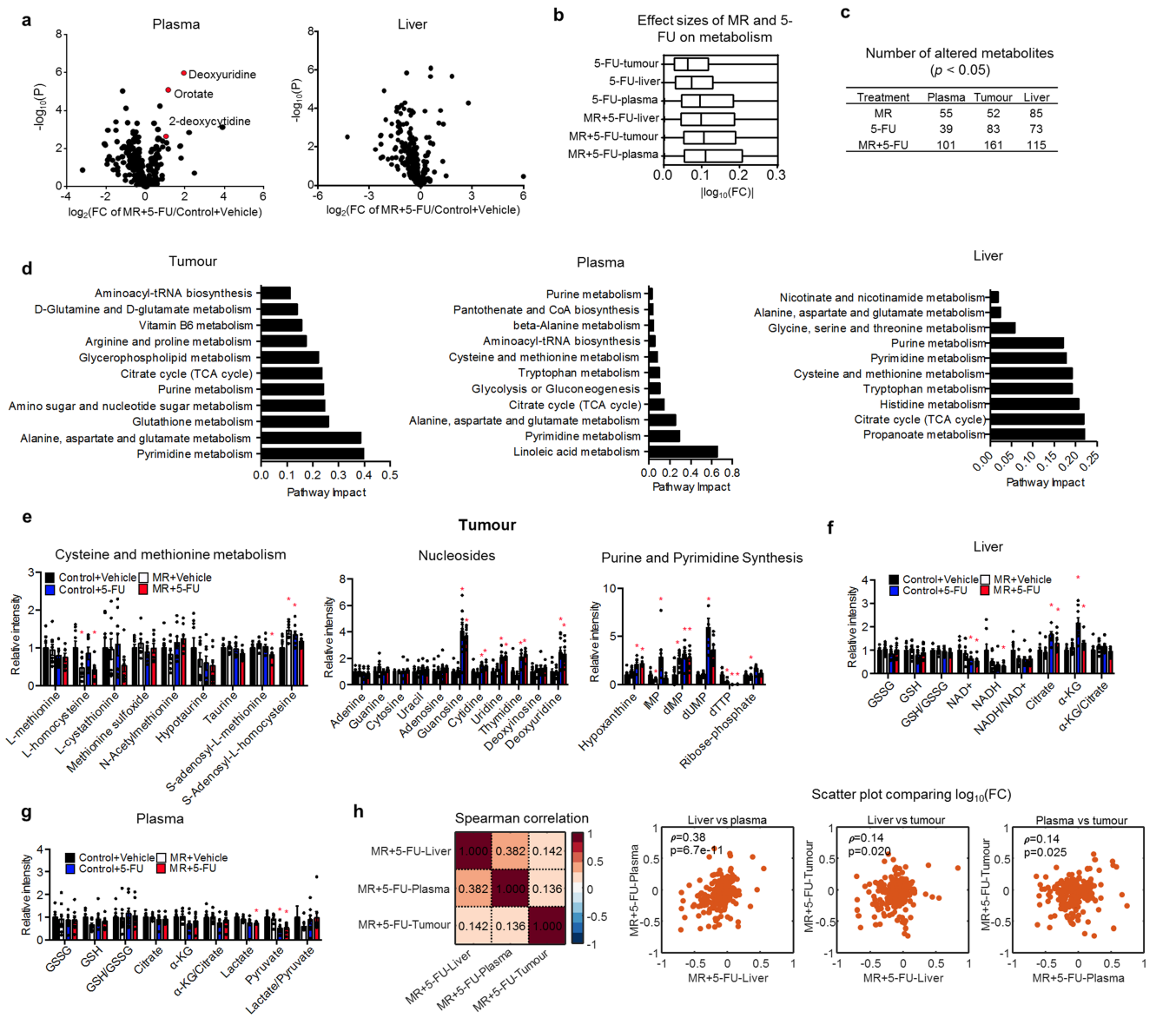
CRC240

Tissue	Methionine related, altered/total	Methionine unrelated, altered/total	p-value by One-sided Fisher's exact test
Plasma	5/17	30/216	0.09
Tumour	5/17	32/216	0.11
Liver	3/17	52/216	0.81

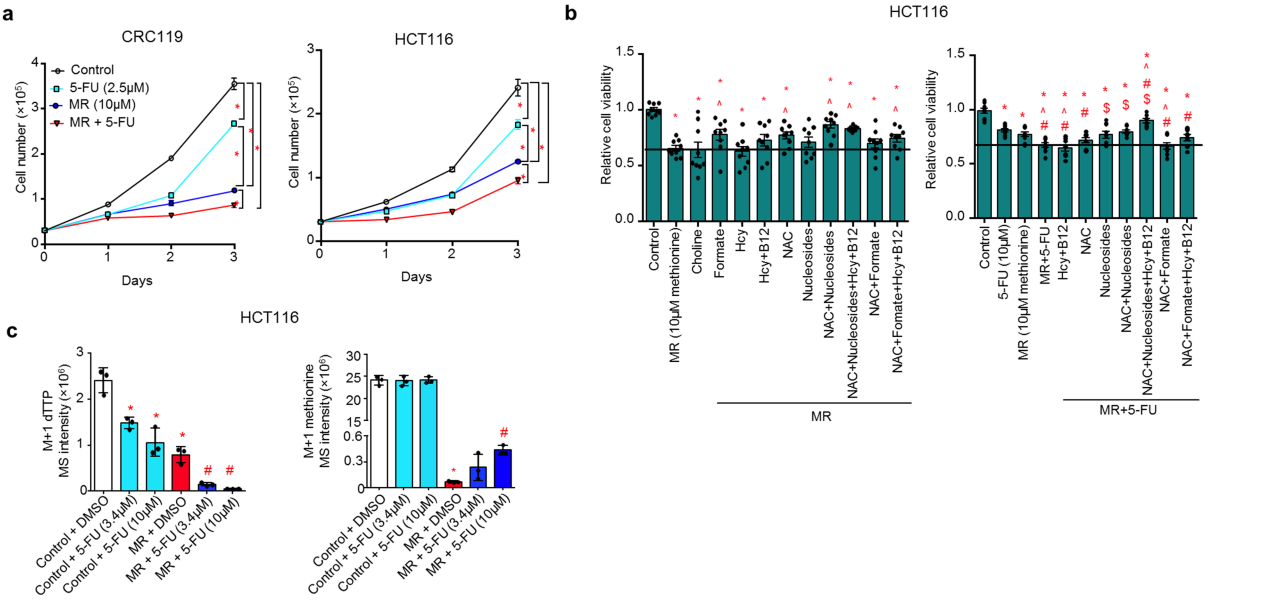
Extended Data Fig. 3 | Methionine restriction leads to specific cell-intrinsic metabolic alterations in tumours. To determine whether the effect of methionine restriction on tumour growth is systemic, cell autonomous or both, we conducted an integrated analysis of global changes in the metabolic network across tumour, plasma and liver within each model from the prevention study shown in Fig. 1f. $n = 8$ mice per group (4 female and 4 male). **a**, Spearman's rank correlation coefficients of fold change of metabolites in tumour, plasma and liver induced by methionine restriction exhibited strong correlations between each tissue pair, with the highest correlation between the tumour and the plasma in both CRC119 and CRC240. **b**, Multidimensional scaling analysis of metabolite fold change in response to methionine restriction. In both models, the fold change of metabolites in tumours showed a higher similarity with those in the plasma than with those in the liver. **c**, **d**, Liver was the most-affected tissue in both models. **c**, The effect of methionine restriction on metabolism in tumour, plasma and liver, evaluated by taking the $\log_{10}$ of the fold change. Box limits are the 25th and 75th percentiles, centre line is the median, and the whiskers are the minimal and maximal values. **d**, Numbers of metabolites significantly altered (*$P < 0.05$, two-tailed Student's t-test) by methionine restriction. $n = 8$ mice per group (4 female and 4 male). **e**, Schematic defining methionine-related metabolites (metabolized from or to methionine within four reaction steps) and methionine-unrelated metabolites. **f**, **g**, A higher proportion of altered metabolites was methionine-related in plasma and tumour compared to liver, in which metabolites altered by methionine restriction were nearly equally distributed between methionine-related and methionine-unrelated groups. **f**, Fraction of significantly (*$P < 0.05$, two-tailed Student's t-test) altered metabolites for methionine-related and methionine-unrelated metabolites in tumour, liver and plasma. **g**, Numbers of total and significantly altered metabolites for methionine-related and methionine-unrelated metabolites in tumour, liver and plasma. P values were determined by one-sided Fisher's exact test.



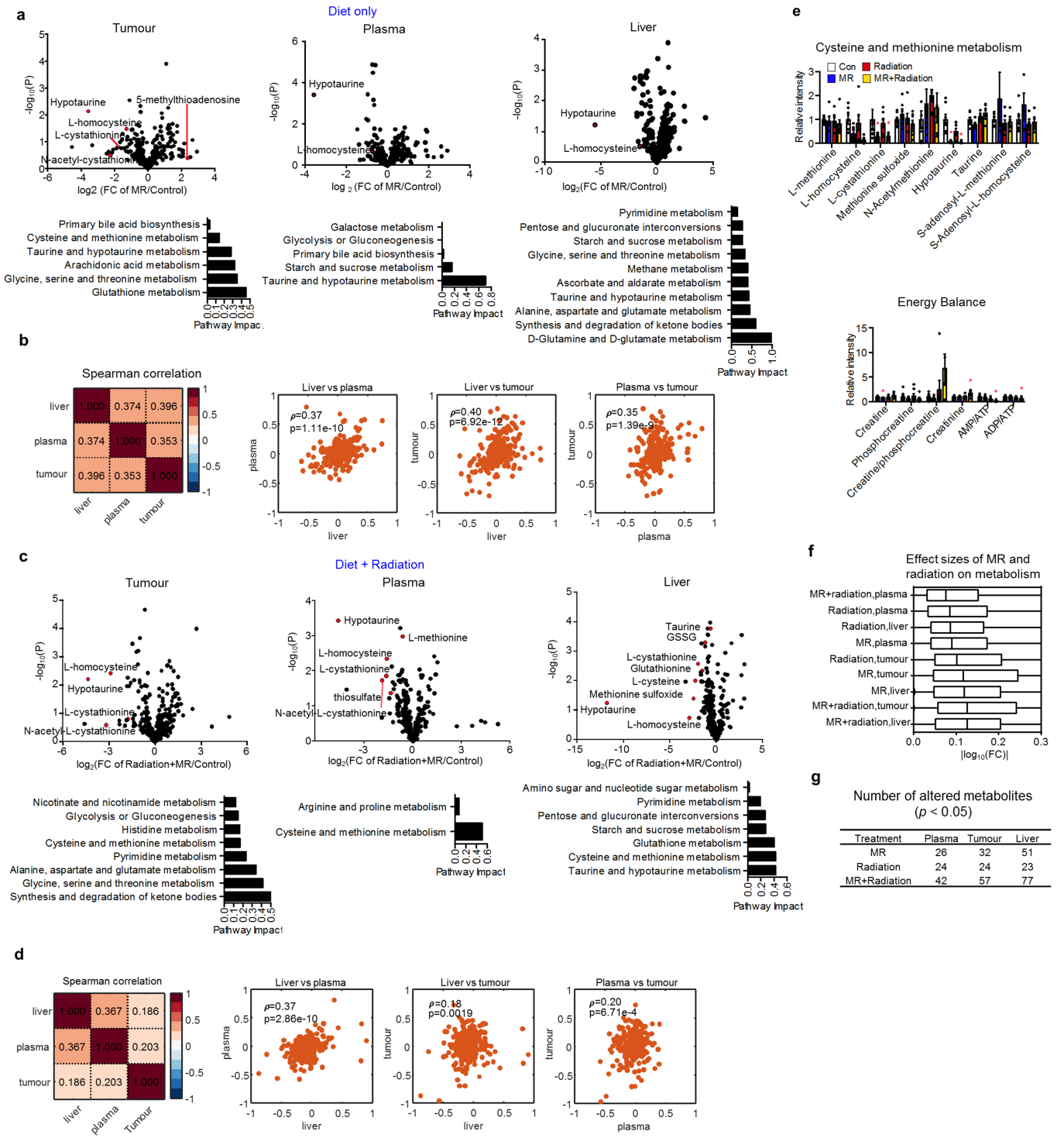
Extended Data Fig. 4 | Methionine restriction inhibits cell proliferation, and most substantially alters cysteine and methionine metabolism in primary colorectal cancer cells. a, Relative cell numbers in CRC119 and CRC240 primary tumour cells treated with different doses of methionine for 72 h. Mean ± s.d. $n = 3$ biologically independent samples; similar results were obtained from 3 independent experiments. *$P < 0.05$, two-tailed Student's t-test. **b**, Volcano plots of metabolites in cells cultured in 0 or 100 μM methionine for 24 h. P values were determined by two-tailed Student's t-test. **c**, Left, Venn diagram of significantly changed (*$P < 0.05$, two-tailed Student's t-test) metabolites in CRC119 and CRC240 primary cells cultured with no methionine versus control (100 μM methionine), and pathway analysis of metabolites that were commonly changed. Right, fold change of metabolites in cysteine and methionine metabolism, and pyrimidine metabolism, in CRC119 and CRC240 primary cells treated with 0 or 100 μM methionine. Mean ± s.d. $n = 3$ biologically independent samples. *$P < 0.05$, two-tailed Student's t-test. **d**, Relative fold change of intensity of amino acids by methionine deprivation in CRC119 and CRC240 primary cells. Mean ± s.d. $n = 3$ biologically independent samples. *$P < 0.05$, two-tailed Student's t-test.



Extended Data Fig. 5 | Dietary restriction of methionine sensitizes PDX models of colorectal cancer to 5-FU chemotherapy. a, Volcano plots of metabolites in plasma and liver altered by the combination of dietary restriction of methionine and 5-FU treatment. *P* values were determined by two-tailed Student's *t*-test. **b**, Effect of 5-FU treatment alone, and a combination of methionine restriction and 5-FU treatment, on metabolites in tumour, plasma and liver, evaluated by taking the $\log_{10}$ of the fold change. Box limits are the 25th and 75th percentiles, centre line is the median, and the whiskers are the minimal and maximal values. The data represents metabolites in liver (337), plasma (282) and tumour (332) from $n = 8$ mice per group. **c**, Numbers of metabolites significantly changed by methionine restriction, 5-FU treatment or the combination of methionine restriction and 5-FU treatment in plasma, tumour and liver. $^*P < 0.05$, two-tailed Student's *t*-test. **d**, Pathway analysis of metabolites significantly changed ($^*P < 0.05$, two-tailed Student's *t*-test) by methionine restriction, 5-FU treatment or the combination of dietary methionine restriction and 5-FU treatment (false discovery rate < 0.5). **e**–**g**, Relative intensity of metabolites related to cysteine and methionine metabolism, and nucleotide metabolism, in tumour (**e**) and redox balance in liver (**f**) and plasma (**g**). Mean ± s.e.m. $n = 8$ mice per group. $^*P < 0.05$, two-tailed Student's *t*-test. **h**, Spearman's rank correlation coefficients of methionine restriction and 5-FU-induced fold change of metabolites in tumour, plasma and liver from mice on dietary methionine restriction and with 5-FU treatment.

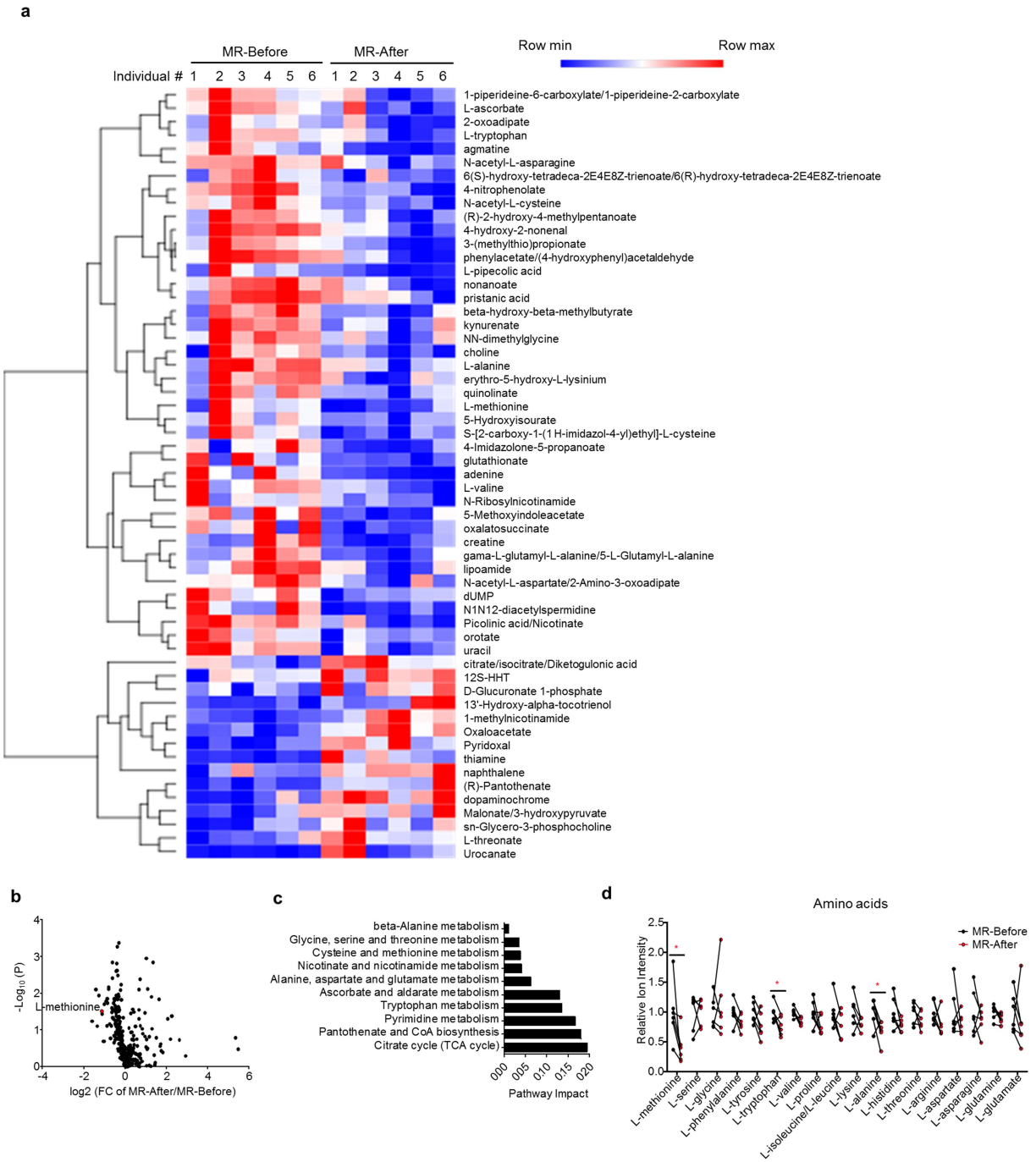


Extended Data Fig. 6 | Inhibition of cell growth mediated by methionine restriction is largely due to interruptions to the production of nucleosides and redox balance. a, The synergic effects of methionine restriction and 5-FU treatment in CRC119 primary cells and HCT116 cells were evaluated by cell counting. Mean ± s.e.m. $n = 3$ biological replicates. $^*P < 0.05$ by two-tailed Student's *t*-test. **b**, The rescue effect of choline, formate, homocysteine, homocysteine with vitamin B12, nucleosides and NAC, alone or in combination, on the inhibition of HCT116 cell proliferation mediated by methionine restriction. Mean ± s.e.m., $n = 9$ biologically independent samples from 3 independent experiments. $^*P < 0.05$ versus control, $^P < 0.05$ versus methionine restriction, $\#P < 0.05$ versus 5-FU treatment, $\$P < 0.05$ versus methionine restriction + 5-FU treatment, by two-tailed Student's *t*-test. **c**, Mass intensity for [M + 1] dTTP and [M + 1] methionine in HCT116 cells from the experiment described in Fig. 2h. Mean ± s.d. $n = 3$ biologically independent samples. $^*P < 0.05$ versus control and $\#P < 0.05$ versus methionine restriction, by two-tailed Student's *t*-test.

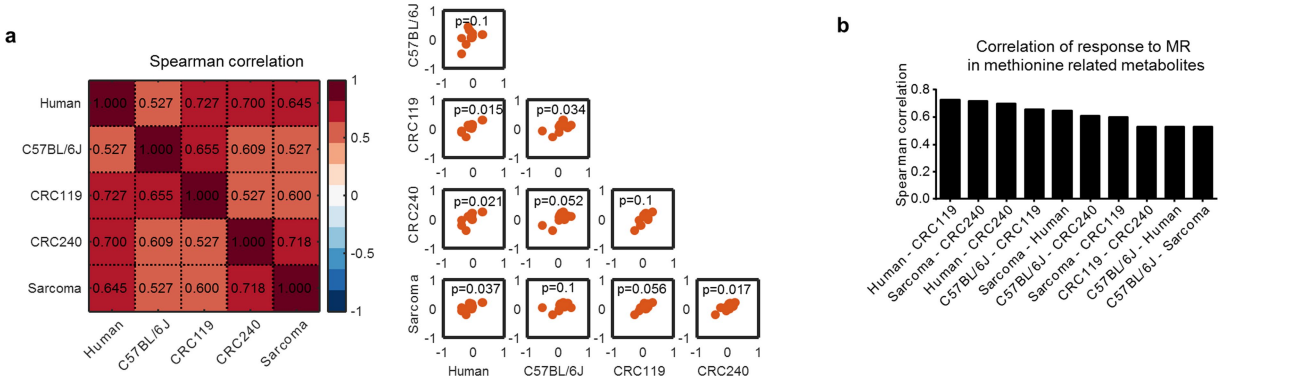


Treatment	Plasma	Tumour	Liver
MR	26	32	51
Radiation	24	24	23
MR+Radiation	42	57	77

Extended Data Fig. 7 | Dietary restriction of methionine sensitizes mouse models of RAS-driven autochthonous sarcoma to radiation. **a**, Volcano plots of metabolites in tumour, plasma and liver, and pathway analysis of metabolites significantly changed (*$P < 0.05$, two-tailed Student's *t*-test) by dietary restriction of methionine alone (false discovery rate < 1). **b**, Spearman's rank correlation coefficients of fold change of metabolites in tumour, plasma and liver induced by methionine restriction. **c**, Volcano plots of metabolites in tumour, plasma and liver, and pathway analysis of metabolites significantly changed (*$P < 0.05$, two-tailed Student's *t*-test) by dietary restriction of methionine and radiation (false discovery rate < 0.5). **d**, Spearman's rank correlation coefficients of fold change of metabolites in tumour, plasma and liver induced by methionine restriction and radiation. **e**, Relative intensity of metabolites related to cysteine and methionine metabolism, and energy balance in tumours. Mean ± s.d. $n = 7$ mice per group, except for the methionine-restriction group ($n = 6$). *$P < 0.05$ versus control, by two-tailed Student's *t*-test. **f**, **g**, The largest effects on metabolism occurred in the combination of diet and radiation. **f**, Effect of methionine restriction and radiation alone, or in combination, on metabolites in tumour, plasma and liver, evaluated by taking the $\log_{10}$ of fold change. Box limits are the 25th and 75th percentiles, centre line is the median, and the whiskers are the minimal and maximal values. The data represent metabolites in liver (319), plasma (308) and tumour (332) from $n = 7$ mice per group, except for the methionine-restriction group ($n = 6$). **g**, Numbers of metabolites significantly changed (*$P < 0.05$, two-tailed Student's *t*-test) by methionine restriction and radiation alone, or in combination.



Extended Data Fig. 8 | Dietary restriction of methionine can be achieved in humans. a, Heat map of significantly changed (*$P < 0.05$, two-tailed Student's *t*-test) plasma metabolites by dietary intervention, in six human subjects. **b**, Volcano plot of plasma metabolites. *P* values were determined by two-tailed Student's *t*-test. **c**, Pathway analysis of altered (*$P < 0.05$, two-tailed Student's *t*-test) plasma metabolites. **d**, Relative intensity of amino acids in plasma. $n = 6$ biologically independent humans. *$P < 0.05$ by two-tailed Student's *t*-test.



Extended Data Fig. 9 | Comparative metabolic effects of methionine restriction across mouse models and humans. a, Spearman's rank correlation coefficients of fold changes of methionine-related metabolites induced by methionine restriction (defined in Extended Data Fig. 3f) in plasma samples from non-tumour bearing C57BL/6J mice, CRC119 and CRC240 mouse models, the sarcoma mouse model and healthy human subjects. **b**, Spearman's rank correlation coefficients among different models in **a**, ranked from the highest to the lowest.

Extended Data Table 1 | Methionine concentrations in plasma, tumour and liver across mouse models and humans

Model	Tissue	Control diet	MR diet
CRC119	Plasma	37.58 ± 6.08	27.44 ± 4.28*
	Tumour	31.23 ± 9.10	20.08 ± 2.36*
	Liver	29.99 ± 2.38	18.68 ± 4.12*
CRC240	Plasma	52.35 ± 12.19	32.98 ± 3.60*
	Tumour	70.77 ± 25.73	31.82 ± 3.05*
	Liver	18.97 ± 3.10	12.393 ± 1.86*
CRC119 (Vehicle)	Plasma	35.93 ± 6.02	33.88 ± 7.97
	Tumour	33.52 ± 10.31	31.50 ± 12.12
	Liver	17.37 ± 4.18	16.68 ± 3.88
CRC119 (5-FU)	Plasma	32.13 ± 5.29	31.31 ± 11.55
	Tumour	26.70 ± 7.15	29.45 ± 11.04
	Liver	16.50 ± 5.14	16.61 ± 2.39
Sarcoma	Plasma	37.75 ± 5.27	35.86 ± 14.57
	Tumour	46.84 ± 10.01	38.39 ± 10.32
	Liver	10.44 ± 4.74	10.45 ± 3.44
Sarcoma (Radiation)	Plasma	24.44 ± 7.56	25.87 ± 4.11#
	Tumour	43.10 ± 11.86	37.22 ± 4.97#
	Liver	10.43 ± 2.98	6.55 ± 1.9*
C57BL/6J	Plasma	99.89 ± 16.90	28.79 ± 4.59*
Human	Plasma	13.74 ± 5.19	6.55 ± 4.02*

Tissues were collected at the end of each experiment. Concentrations in tissues were estimated in μM, assuming that 1 g wet tissue weight = 1 ml. Quantification was performed by using ^{13}C-labelled standards for each amino acid, which were added before extraction. Cyclic loess normalization and linear regression were applied in quantification of methionine in samples without ^{13}C-labelled standards. Values are mean ± s.d. $n = 8$ for CRC119, CRC240, CRC119 (vehicle) and CRC119 (5-FU), 7 for sarcoma on the control diet, 6 for sarcoma on the methionine-restricted diet, 7 for sarcoma (radiation), 5 for C57BL/6J mice and 6 for humans. *$P < 0.05$, by two tailed Student's *t*-test between the control diet and the methionine-restricted diet; #$P < 0.05$, by two tailed Student's *t*-test between the control diet and the methionine-restricted diet + radiation.

nature research

Corresponding author(s): Jason Locasale

Last updated by author(s): May 21, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No specific software was used for data collection.
Data analysis	Raw metabolomics data were processed with Sieve 2.0. from Thermo Scientific, Pathway analysis of metabolites was carried out with software Metaboanalyst (http://www.metaboanalyst.ca/MetaboAnalyst/) using the KEGG pathway database (http://www.genome.jp/kegg/). Cross-tissue comparison of metabolite profiles in PDX and sarcoma models were carried out in MATLAB R2018b. GraphPad PRISM 6.0 and Microsoft Excel were used for statistical analysis.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The metabolomics data reported in this study has been deposited to Mendeley Data (DOI: doi:10.17632/zs269d9fvb.1). Source code has been deposited to GitHub (https://github.com/LocasaleLab/Dietary_methionine_restriction). Source data are provided for all figures.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes were given in the manuscript. Efficacy studies in mice were conducted with 5 or more mice per group. No statistical methods were used to predetermine sample size.
Data exclusions	No data were excluded.
Replication	Replication numbers were reported in the text or in the methodology sections. Where appropriate a measure of the error was reported. All metabolic tracing and profiling, and cell counting study, were done once. For MTT assay, experiments were repeated at least three times with similar results. Biological replicates were produced.
Randomization	For all animal studies, mice were randomized into different research groups.
Blinding	For the dietary studies, the same investigators carried out the dietary treatment and downstream and analysis, so were not blinded to group allocation.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	CRC119 and CRC240 cell lines were developed from their respective CRC PDXs in Dr. David Hsu's lab at Duke University. HCT116 cell line was a gift from Dr. Lewis Cantley's laboratory.
Authentication	Cell lines were authenticated at the Duke University DNA Analysis Facility by analyzing DNA samples from each cell lines for polymorphic short tandem repeat (STR) markers using the GenePrint 10 kit from Promega (Madison, WI, USA).
Mycoplasma contamination	All cell lines were negative for mycoplasma contamination
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	The following mouse strains were used in the manuscript: C57BL/6J mice: male, 12-week-old; NOD.CB17-PrkdcSCID-J mice: male and female, 8-10-week-old; NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ mice: male and female, 8-10-week-old; C57BL/6J × 129SvJ mice carrying p53FRT and FSF-KrasG12D: male and female, 6-10-week-old;

Wild animals	No wild animals were involved in this study.
Field-collected samples	The study did not involve samples collected from field.
Ethics oversight	All animal procedures and studies were approved by the Institutional Animal Care and Use Committee (IACUC) at Duke University. All animal experiments were performed in accordance with relevant guidelines and regulations.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics	healthy human subjects (both male and female) recruited in the dietary study aged between 49-58 years old
Recruitment	Healthy adults of mixed gender were recruited by fliers and word of mouth and, as assessed for initial eligibility by telephone interview, were free of disease and currently not taking certain medications including anti-inflammatory drugs, corticosteroids, statins, thyroid drugs, and oral contraceptives. Final eligibility was assessed by standard clinical chemistry and hematology analyses. Written consent was obtained from eligible subjects. We are unaware of any potential self-selection bias or other biases present.
Ethics oversight	The controlled feeding study in healthy humans was conducted at Penn State University Clinical Research Center (CRC) and approved by the Institutional Review Board of the Penn State College of Medicine in accordance with the Helsinki Declaration of 1975 as revised in 1983 (IRB# 32378). We have complied with all relevant ethical regulations.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

LETTER

Corrected: Publisher Correction

https://doi.org/10.1038/s41586-019-1426-6

Intercellular interaction dictates cancer cell ferroptosis via NF2–YAP signalling

Jiao Wu[1,2,6], Alexander M. Minikes[2,3,6], Minghui Gao[2,4], Huijie Bian[1], Yong Li[1], Brent R. Stockwell[5], Zhi-Nan Chen[1]* & Xuejun Jiang[2]*

Ferroptosis, a cell death process driven by cellular metabolism and iron-dependent lipid peroxidation, has been implicated in diseases such as ischaemic organ damage and cancer[1,2]. The enzyme glutathione peroxidase 4 (GPX4) is a central regulator of ferroptosis, and protects cells by neutralizing lipid peroxides, which are by-products of cellular metabolism. The direct inhibition of GPX4, or indirect inhibition by depletion of its substrate glutathione or the building blocks of glutathione (such as cysteine), can trigger ferroptosis[3]. Ferroptosis contributes to the antitumour function of several tumour suppressors such as p53, BAP1 and fumarase[4–7]. Counterintuitively, mesenchymal cancer cells—which are prone to metastasis, and often resistant to various treatments—are highly susceptible to ferroptosis[8,9]. Here we show that ferroptosis can be regulated non-cell-autonomously by cadherin-mediated intercellular interactions. In epithelial cells, such interactions mediated by E-cadherin suppress ferroptosis by activating the intracellular NF2 (also known as merlin) and Hippo signalling pathway. Antagonizing this signalling axis allows the proto-oncogenic transcriptional co-activator YAP to promote ferroptosis by upregulating several ferroptosis modulators, including ACSL4 and TFRC. This finding provides mechanistic insights into the observations that cancer cells with mesenchymal or metastatic property are highly sensitive to ferroptosis[8]. Notably, a similar mechanism also modulates ferroptosis in some non-epithelial cells. Finally, genetic inactivation of the tumour suppressor NF2, a frequent tumorigenic event in mesothelioma[10,11], rendered cancer cells more sensitive to ferroptosis in an orthotopic mouse model of malignant mesothelioma. Our results demonstrate the role of intercellular interactions and intracellular NF2–YAP signalling in dictating ferroptotic death, and also suggest that malignant mutations in NF2–YAP signalling could predict the responsiveness of cancer cells to future ferroptosis-inducing therapies.

Cellular metabolism has a crucial role in ferroptosis[1,2]. To study the underlying mechanisms further, we manipulated cellular metabolism by altering the ingredients of culture medium or cell number in culture. Unexpectedly, we observed that cells became more resistant to ferroptosis when approaching high confluence. In HCT116 human colon cancer cells, higher cell confluence conferred resistance to ferroptosis and associated lipid peroxidation, induced by cystine starvation, the cystine transporter inhibitor erastin and the GPX4 inhibitor RSL3 (Fig. 1a, b and Extended Data Fig. 1a–e). Using corresponding pharmacological inhibitors, we confirmed that cells underwent ferroptosis rather than apoptosis or necroptosis under these conditions (Extended Data Fig. 1f, g). Notably, previous published observations also suggest cell-density-dependent ferroptosis: GPX4-null mouse embryonic fibroblasts (MEFs) were able to grow when seeded at high density or as 3D spheroids, but died rapidly after passage at low density[12,13].

To examine whether such dependence on cell density is a general property of ferroptosis, we tested a panel of human epithelial cancer cell lines (Fig. 1c). Most tested cell lines showed cell density dependence, with two exceptions: MDA-MB-231 (MDA231) cells were always sensitive to ferroptosis, whereas BT474 cells were always resistant, regardless of density. To better mimic the in vivo context, we cultured these cells into 3D tumour spheroids. Consistently, erastin triggered more prominent cell death in spheroids formed by MDA231 and H1650 cells (Fig. 1d, e). A possible explanation for this phenomenon is that high cell density more rapidly depletes glutamine (required for cysteine-deprivation-induced ferroptosis[4,14]). However, replenishing glutamine to confluent cells did not restore cell death (Extended Data Fig. 1h).

Cells tend to forge cell–cell contacts with higher cell confluence, and E-cadherin (ECAD) is an important mediator of intercellular contact in epithelial cells[15]. Expression of ECAD correlated with sensitivity to ferroptosis: ECAD was undetectable in MDA231 cells and very low in H1650 cells (Fig. 1f). As cell density increased, ECAD expression increased and became enriched at sites of cell–cell contact in cells that underwent density-dependent ferroptosis; BT474 cells, which are resistant to ferroptosis regardless of confluence, expressed high levels of ECAD even at low cell density (Extended Data Fig. 2a–d). Strong expression of ECAD was detected in spheroids generated from HCT116 cells, but not in those generated from MDA231 cells (Extended Data Fig. 2e). To determine further whether ECAD has a causative role, we tested whether inhibition of ECAD dimerization would sensitize confluent cells to ferroptosis. Indeed, an anti-ECAD antibody that blocks its intercellular dimerization markedly increased the sensitivity of confluent cells to ferroptosis (Extended Data Fig. 2f). ECAD depletion (ΔECAD) rendered confluent HCT116 cells sensitive to ferroptosis (Extended Data Fig. 2g–i). ECAD depletion did not induce expression of N-cadherin (NCAD) in HCT116 cells (Extended Data Fig. 2g). Re-expression of full-length ECAD, but not a truncated mutant lacking the ectodomain (required for intercellular dimerization of ECAD), restored resistance to ferroptosis in ΔECAD cells (Fig. 1g, h and Extended Data Fig. 2j, k).

ECAD-mediated intercellular interaction can signal to the Hippo pathway[16,17], which regulates a plethora of biological events that includes control of proliferation and of organ size[18,19]. The Hippo pathway involves the tumour suppressor NF2 and a kinase cascade consisting of MST1, MST2, LATS1 and LATS2. NF2 has been shown to activate the Hippo signalling pathway by inhibiting CRL4–DCAF1, a ubiquitin ligase complex that promotes proteasomal degradation of LATS1 or LATS2[20,21]. LATS1 and LATS2 phosphorylate the pro-oncogenic transcription co-activator YAP, leading to its nuclear exclusion and inactivation. As expected, as HCT116 cells grew more confluent, increased phosphorylation and decreased nuclear localization of YAP were observed (Extended Data Fig. 3a, b); ECAD knockout or *NF2* RNA interference (RNAi) diminished cell-density-regulated nuclear exclusion of YAP (Extended Data Fig. 3c–g, Supplementary Table 1). To confirm further that YAP is functionally activated under these conditions, we used an 8xGTIIC-luciferase reporter assay that monitors

[1]National Translational Science Center for Molecular Medicine, Department of Cell Biology, School of Basic Medicine, Air Force Medical University, Xi'an, China. [2]Cell Biology Program, Memorial Sloan-Kettering Cancer Center, New York, NY, USA. [3]BCMB Allied Program, Weill Cornell Graduate School of Medical Sciences, New York, NY, USA. [4]The HIT Center for Life Sciences, School of Life Science and Technology, Harbin Institute of Technology, Harbin, China. [5]Department of Biological Sciences, Department of Chemistry, Columbia University, New York, NY, USA. [6]These authors contributed equally: Jiao Wu, Alexander M. Minikes. *e-mail: znchen@fmmu.edu.cn; jiangx@mskcc.org

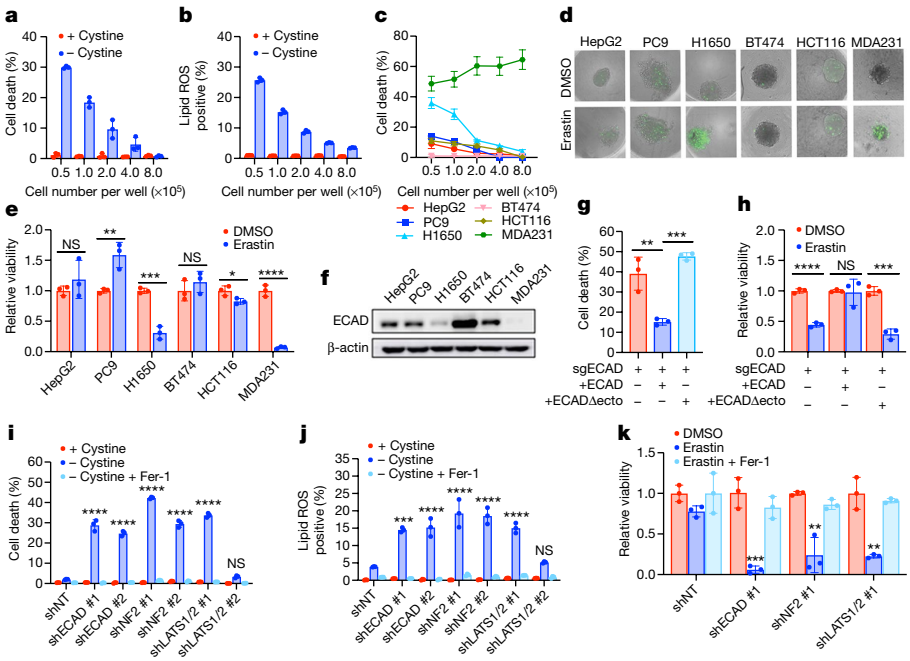


Fig. 1 | E-cadherin and the Hippo pathway regulate ferroptosis in a manner dependent on cell density. a, **b**, HCT116 cells were seeded at the indicated density in 6-well plates and cultured for 24 h. **a**, Ferroptosis was measured after cystine starvation for 30 h, by SYTOX Green staining followed by flow cytometry. **b**, Production of lipid reactive oxygen species (ROS) was measured and quantified after cystine starvation for 24 h, by staining of C11-BODIPY followed by flow cytometry. **c**, Ferroptosis was measured in the 6 indicated cell lines after cystine starvation for 30 h. **d**, **e**, Spheroids generated from the indicated cell lines were cultured for 72 h and treated with 15 μM erastin for 30 h. Dead cells were stained by SYTOX Green (**d**) (original magnification, ×40) and cell viability was assayed by measuring cellular ATP levels (**e**). NS, not significant ($P = 0.3757$, 0.3572; from left to right). $^{*}P = 0.0323$, $^{**}P = 0.0086$, $^{***}P = 0.0004$, $^{****}P < 0.0001$; two-tailed *t*-test. **f**, Western blot of ECAD levels in the indicated cell lines. Image represents three experiments; see Supplementary Fig. 1 for raw blots. **g**, Ferroptosis after cystine starvation for 30 h in HCT116 cells depleted of ECAD using single-guide RNA (sgECAD) and in ECAD-depleted cells expressing full-length or ectodomain-truncated (Δecto) ECAD. $^{**}P = 0.0025$, $^{***}P = 0.0005$; one-way analysis of variance (ANOVA). **h**, Viability of spheroids generated from cells as in **g** after treatment with erastin or dimethylsulfoxide (DMSO) control. NS, $P = 0.8683$. $^{***}P = 0.0004$, $^{****}P < 0.0001$; two-tailed *t*-test. **i**, Ferroptosis of HCT116 cells after cystine starvation for 30 h and the addition of 2 μM ferrostatin-1 (Fer-1). NS, $P = 0.6880$. $^{****}P < 0.0001$; one-way ANOVA. Note that shLATS1/2 #2 did not knockdown LATS2 (Extended Data Fig. 3e) and thus did not sensitize cells to ferroptosis. **j**, Lipid ROS production of cells as in **i**. NS, $P = 0.9383$. $^{***}P = 0.0001$, $^{****}P < 0.0001$; one-way ANOVA. **k**, Viability of spheroids generated from HCT116 cells after the indicated treatments. $^{**}P = 0.0012$, 0.0010 (left to right), $^{***}P = 0.0002$; one-way ANOVA. All data are mean ± s.d. from $n = 3$ biological replicates.

the transcriptional activity of YAP (also known as YY1AP1) with its primary binding partners, the TEAD family of transcription factors[22]. Low cell density, loss of ECAD or *NF2* RNAi all increased YAP activity and upregulated transcription of the canonical YAP targets *CTGF* and *CYR61* (Extended Data Fig. 3h–l). Knockdown of *ECAD*, *NF2*, *LATS1* and *LATS2* all sensitized HCT116 cells to ferroptosis in cell culture and spheroids (Fig. 1i–k and Extended Data Fig. 4a–c). Notably, knockdown of *ECAD*, *NF2*, *LATS1* and *LATS2* did not decrease cell proliferation within the time frame of the experiment, ruling out the possibility that increased ferroptosis was due to reduced cell confluence (Extended Data Fig. 4d). In addition, p21-activated kinase (PAK) can phosphorylate and inactivate NF2[17]. Consistently, constitutively active PAK (PAK-CAAX), but not its inactive mutant form (PAK-CAAX(K290R)), enhanced YAP activity and ferroptosis (Extended Data Fig. 4e–h). Together, ECAD and Hippo signalling negatively regulate ferroptosis.

Heterozygous deletion and loss-of-function mutations of the *NF2* gene are detected with high frequency in malignant mesothelioma, and inactivation of either NF2 or LATS1 or LATS2 is observed in approximately 50% of patients with malignant mesothelioma[10,11]. We assessed NF2 status and ferroptosis sensitivity in a cohort of human malignant mesothelioma cell lines. Of ten patient-derived cell lines we examined, four had wild-type NF2 expression, and six were NF2-defective[21] (Fig. 2a). All NF2 wild-type cells expressed a cadherin protein (not necessarily ECAD) and either LATS1 or LATS2 (Fig. 2a and Extended Data Fig. 5a). Several NF2-mutant cell lines can undergo potent ferroptosis even at the highest tested density and in spheroids, whereas all NF2 wild-type cells were relatively insensitive to ferroptosis under the same conditions (Fig. 2b, c and Extended Data Fig. 5b). Consistently, *NF2* RNAi sensitized confluent NF2 wild-type 211H cells to ferroptosis (Fig. 2d, e and Extended Data Fig. 5c, d), and NF2 reconstitution in confluent, NF2-defective Meso33 cells decreased nuclear localization of YAP and mitigated ferroptosis (Extended Data Fig. 5e–h). Furthermore, we generated a doxycycline (Dox)-inducible system to express NF2 in Meso33 cells (Fig. 2f). Indeed, Dox-induced restoration of NF2 inhibited ferroptosis at high density and in a spheroid model (Fig. 2g, h and Extended Data Fig. 5i).

Of the NF2 wild-type mesothelioma cells tested, only H-meso cells expressed ECAD (Fig. 2a). 211H cells express NCAD in a manner dependent on cell density (Extended Data Fig. 6a). We found that NCAD was similarly able to suppress ferroptosis in these cells and signal through the NF2–YAP axis (Extended Data Fig. 6b–k). We also observed cell-density-dependent, NF2-regulated ferroptosis in MEFs, which are not of epithelial origin (Extended Data Fig. 7a–k). Notably, we also observed a modest effect of cell density in a Burkitt lymphoma cell line, which does not express YAP or its homologue TAZ (Extended Data Fig. 7l–m), suggesting an alternative mechanism (cystine production by transsulfuration could be a contributor, as previously reported[23]).

The correlation between YAP activity and ECAD- or NF2-regulated ferroptosis prompted us to perform additional functional experiments to determine whether YAP promotes ferroptosis. The YAP(S127A) mutant cannot be phosphorylated by LATS1 or LATS2 at the Ser127 residue, thus enhancing nuclear retention and transcriptional

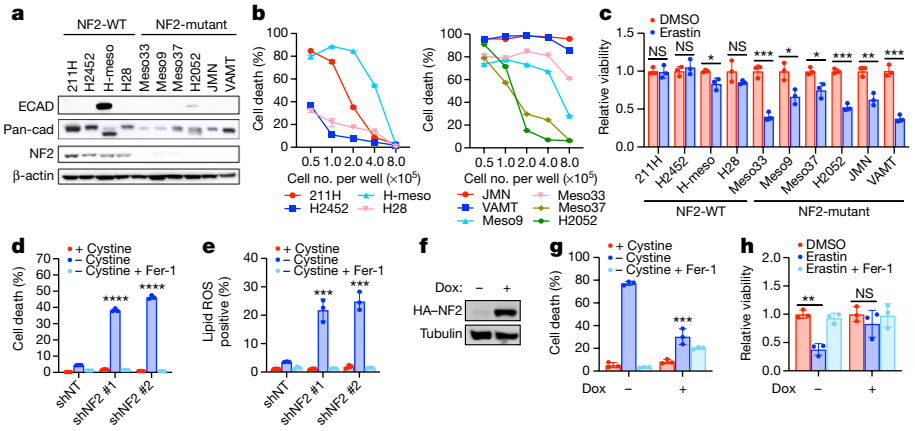


Fig. 2 | NF2 mediates cell-density-dependent inhibition of ferroptosis in mesothelioma cells. a, Western blot analysis of the expression of ECAD, pan-cadherin (pan-cad) and NF2 in a panel of mesothelioma cell lines cultured at high confluence. **b**, NF2 wild-type (WT; left) or mutant (right) mesothelioma cells were seeded at the indicated densities, and cell death was measured after cystine starvation for 24 h. **c**, Spheroids generated from the indicated cell lines were treated with 10 μM erastin for 24 h before the measurement of cell viability by ATP levels. NS, $P = 0.8860$, 0.4981, 0.1474 (left to right). $^{*}P = 0.0203$, 0.0180, 0.0162 (left to right), $^{**}P = 0.0033$, $^{***}P = 0.0005$, 0.0001, 0.0003 (left to right); two-tailed *t*-test. **d**, Cell death was measured in confluent cells after cystine starvation for 24 h, with or without the addition of 2 μM Fer-1. $^{****}P < 0.0001$; one-way ANOVA. **e**, Lipid ROS production of cells as in **d** after 18 h of treatment. $^{***}P = 0.0005$, 0.0002 (left to right); one-way ANOVA. **f**, Western blot analysis confirming expression of NF2 in Meso33 cells containing Dox-inducible NF2 after 48 h of treatment with 1 μg ml^{-1} Dox. HA, haemagglutinin tag. **g**, Cells in the presence or absence of Dox after cystine starvation for 12 h. $^{***}P = 0.0003$; two-tailed *t*-test. **h**, Spheroids were grown in the presence or absence of Dox for 72 h, at which point 10 μM erastin was added. Cell viability was measured by ATP levels after 24 h. NS, $P = 0.3393$. $^{**}P = 0.0010$; two-tailed *t*-test. All data are mean ± s.d. from $n = 3$ biological replicates.

regulatory activity even at high density[20,24,25] (Extended Data Fig. 8a–d). HCT116 or 211H cells that express YAP(S127A) were markedly more sensitive to ferroptosis at high density or in spheroids (Fig. 3a–c and Extended Data Fig. 8e–k). HCT116 cells that lack YAP were no longer sensitized to ferroptosis after *NF2* RNAi (Fig. 3d and Extended Data Fig. 8l), demonstrating that NF2 suppresses ferroptosis by inhibiting YAP activity.

Subsequently, we examined a range of putative YAP and TEAD gene targets that are known regulators of ferroptosis. Putative YAP–TEAD gene targets were selected from the TEAD4 ENCODE chromatin immunoprecipitation followed by high-throughput sequencing (ChIP–seq) datasets GSM1010875 and GSM1010868. Among these genes, we found that transferrin receptor 1 (*TFRC*) and acyl-CoA synthetase long chain family member 4 (*ACSL4*)—both crucial mediators of ferroptosis[14,26]—are genuine targets of the YAP–TEAD complex. Expression of TFRC and ACSL4 decreased with increasing cell density, and TFRC and ACSL4 were both upregulated by depletion of ECAD, knockdown of NF2 or overexpression of YAP(S127A) (Fig. 3e–h). TEAD4 binds

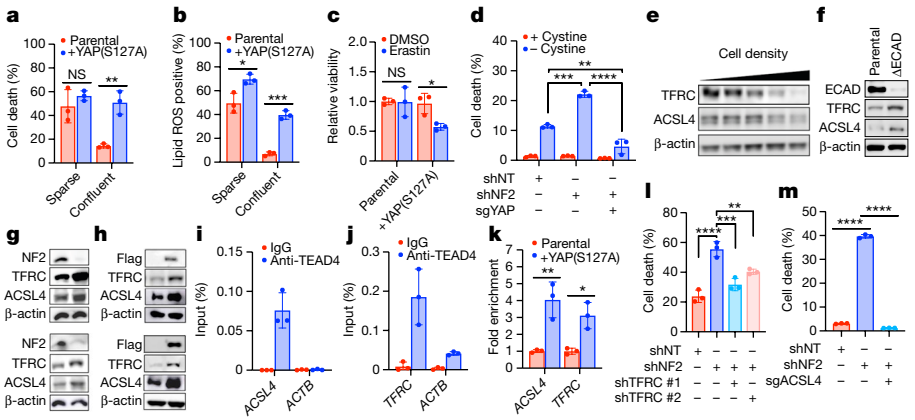


Fig. 3 | The transcriptional regulatory activity of YAP promotes ferroptosis. a, Cells were cultured as indicated. Cell death was measured after cystine starvation for 24 h. NS, $P = 0.3525$. $^{**}P = 0.0031$; two-tailed *t*-test. **b**, Lipid ROS production of cells after cystine starvation for 16 h. $^{*}P = 0.0202$, $^{***}P = 0.0001$; two-tailed *t*-test. **c**, Spheroids generated from parental HCT116 cells and YAP(S127A)-overexpressing cells were treated with erastin or DMSO as indicated, and cell viability was measured by cellular ATP levels. NS, $P = 0.957$. $^{*}P = 0.0200$; two-tailed *t*-test. **d**, Ferroptosis of the indicated cells after cystine starvation for 24 h and transfection with non-targeting shRNA (shNT), *NF2* shRNA (shNF2) or CRISPR–Cas9-mediated knockout of *YAP* (sgYAP). $^{**}P = 0.0043$, $^{***}P = 0.0004$, $^{****}P < 0.0001$; one-way ANOVA. **e**, Western blot analysis of TFRC and ACSL4 in 211H cells seeded at increasing density. **f**, Western blot analysis of TFRC and ACSL4 in parental and ECAD-depleted (ΔECAD) HCT116 cells. **g**, **h**, Western blot analysis of TFRC and ACSL4 in HCT116 (top) or 211H (bottom) cells transfected with shNT or shNF2 (**g**) or overexpressing YAP(S127A) (**h**). **i**, ChIP analysis of TEAD4 binding to the *ACSL4* promoter in 211H cells using control immunoglobulin G (IgG) or an anti-TEAD4 antibody. Values are percentage of input. Quantitative PCR (qPCR) primers were designed based on TEAD4-binding peak regions depicted in the ENCODE TEAD4 ChIP–seq datasets. **j**, TEAD4 binding to the promoter region of *TFRC* was analysed as described in **i**. **k**, ChIP analysis monitoring the occupancy of TEAD4 on the *ACSL4* and *TFRC* promoters in parental or YAP(S127A)-overexpressing 211H cells. Enrichment was calculated based on qPCR relative to the IgG control. $^{*}P = 0.0103$, $^{**}P = 0.0079$; two-tailed *t*-test. **l**, Cell death was measured in HCT116 cells expressing the indicated shRNAs after cystine starvation for 30 h. $^{**}P = 0.0072$, $^{***}P = 0.0004$, $^{****}P < 0.0001$; one-way ANOVA. **m**, Cell death was measured in HCT116 cells expressing indicated shRNA and/or sgRNA, after cystine starvation for 30 h. $^{****}P < 0.0001$; one-way ANOVA. All data are mean ± s.d. from $n = 3$ biological replicates.

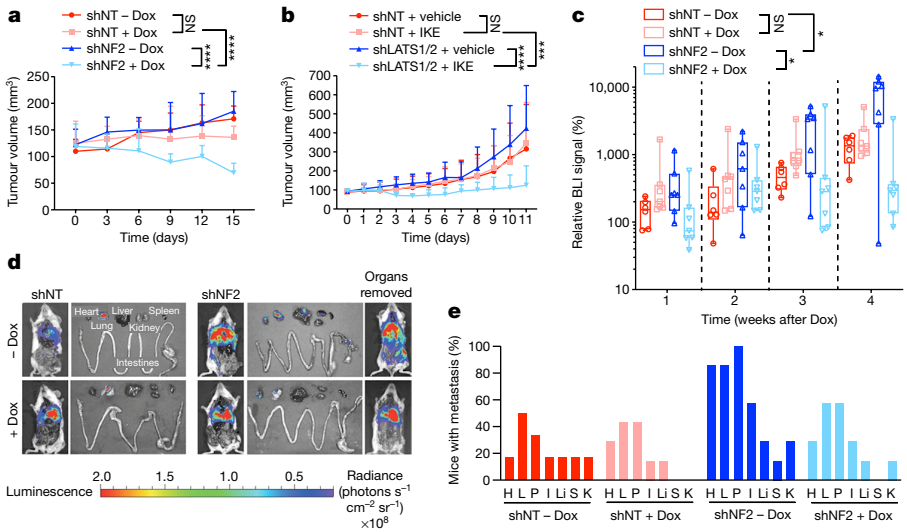


Fig. 4 | NF2 dictates GPX4 dependency in mouse models of mesothelioma. a, Growth curves of tumours derived from GPX4-knockout (GPX4-iKO) 211H cells containing shNF2 or shNT injected subcutaneously into nude mice fed Dox or a normal diet ($n = 8$ per group). NS, $P = 0.6776$. ****$P < 0.0001$; two-way ANOVA. **b**, The indicated HCT116 cells were injected subcutaneously into nude mice ($n = 6$ per group). Tumours were grown to a volume of 90 mm^3, at which point 50 mg kg^{-1} IKE was administered intraperitoneally daily for 12 days. NS, $P = 0.9808$. ***$P = 0.0001$, ****$P < 0.0001$; two-way ANOVA. For knockdown efficiency of LATS1 and LATS2, see Extended Data Fig. 3e. **c**, shNT-GPX4-iKO or shNF2-GPX4-iKO 211H cells were orthotopically injected into the pleural cavity of mice. The percentage change in the relative bioluminescent imaging (BLI) signal (photons s^{-1}) versus time-point 0 is shown. $n = 6$ (shNT − Dox) or 7 mice for each group. Box plots represent median ± interquartile range, whiskers represent the range of values. NS, $P = 0.1545$. *$P = 0.0237$, 0.0287 (top to bottom); two-way ANOVA. **d**, Bioluminescence imaging in excised organs, and in mouse bodies before and after organs were removed. **e**, Percentage of mice in each group with metastases in excised organs. shNT − Dox: $n = 6$; shNT + Dox, shNF2 − Dox, shNF2 − Dox: $n = 7$. H, heart; I, intestines/mesenteric lymph nodes; K, kidneys L, lung; Li, liver; P, peritoneum; S, spleen. Data in **a** and **b** are mean ± s.d.

to the promoter regions of *TFRC* and *ACSL4* genes, and binding was enhanced by overexpression of YAP(S127A) (Fig. 3i–k). Confluent HCT116 cells were sensitized to ferroptosis after the expression of either TFRC or ACSL4, and co-expression of both further enhanced cell death (Extended Data Fig. 8m, n). Conversely, reduced expression of TFRC or ACSL4 mitigated ferroptosis in sensitized cells (Fig. 3l, m and Extended Data Fig. 8o–r). Together, these data indicate that upregulation of TFRC and ACSL4 contributes to the ability of YAP to promote ferroptosis. Notably, co-overexpression of TFRC and ACSL4 did not restore ferroptosis in confluent cells to the level of that in sparse cells, even when the ectopic ACSL4 level was higher than that in sparse cells (Extended Data Fig. 8m, n), which suggests that additional YAP target genes contribute to this process.

As loss of NF2 frequently drives mesothelioma[10,11], we examined whether NF2 status could predict mesothelioma sensitivity to ferroptosis. We generated Dox-inducible, CRISPR–Cas9-mediated GPX4-knockout (GPX4-iKO) 211H cells containing short hairpin RNA (shRNA) against *NF2* (shNF2) or non-targeting shRNA (shNT) (Extended Data Fig. 9a). Spheroids cultured after *NF2* shRNA cells were more sensitive than shNT cells to GPX4 knockout-induced ferroptosis (Extended Data Fig. 9b). We then used shNT-GPX4-iKO cells and shNF2-GPX4-iKO cells to produce subcutaneous xenograft tumours in athymic nude mice. In tumours, knockdown of NF2 increased the expression of TFRC, ACSL4 and nuclear YAP (Extended Data Fig. 9c). The addition of Dox sharply reduced the expression of GPX4 in tumours; in *NF2* RNAi tumours, Dox addition resulted in increased expression of the ferroptosis marker PTGS2[3] and reduced proliferation, as measured by Ki67 staining (Extended Data Fig. 9d). Notably, after Dox addition, shNF2 tumours receded whereas shNT tumours only showed a decrease in growth (Fig. 4a and Extended Data Fig. 9e). Similarly, knockdown of LATS1 and LATS2 rendered xenograft tumours generated by HCT116 cells significantly more sensitive to imidazole ketone erastin (IKE), an erastin derivative amenable for use in vivo[27] (Fig. 4b and Extended Data Fig. 9f).

We next developed an intrapleural mouse model of mesothelioma, by orthotopically implanting shNF2-GPX4-iKO or shNT-GPX4-iKO cells containing a retroviral TK-GFP-luciferase (TGL) reporter. shNF2-GPX4-iKO cells grew more aggressively than shNT-GPX4-iKO cells in mice, consistent with the tumour-suppressive nature of NF2; the addition of Dox reduced the growth of shNF2 tumours, whereas shNT tumours were unaffected (Fig. 4c and Extended Data Fig. 9g). After euthanization, various organs were excised for bioluminescence imaging. shNT tumours grew within the pleural cavity, attaching to the aortic arch, lung or thoracic muscles, whereas shNF2 tumours metastasized to the pericardium, peritoneum, abdominal organs including liver, intestine and distal lymph nodes (Fig. 4d, e)—consistent with previous reports that NF2 loss enhances metastasis of mesothelioma[10]. Supporting this notion, spheroids cultured from *NF2* shRNA cells extended more finger-like protrusions into Matrigel (Extended Data Fig. 9h). Importantly, the metastatic capability of *NF2* shRNA tumours was reduced by Dox-induced knockout of GPX4 (Fig. 4d, e). Therefore, NF2 status might be useful as a biomarker to predict mesothelioma metastasis and responsiveness to the induction of ferroptotic cell death.

Sorafenib, an orally administered multi-kinase inhibitor used for the treatment of hepatocellular carcinoma and renal cell carcinoma, also induces ferroptosis by inhibition of the x_c^- amino acid antiporter[28]. The potential for sorafenib as a therapy for malignant mesothelioma has been tested in clinical trials. The results suggest that sorafenib can stabilize the disease but achieves responses in only a small proportion of unselected patients[29,30]. However, these trials did not examine the genetic status of the NF2–Hippo pathway. We found that sorafenib induced ferroptosis in a manner that is dependent on cell density and Hippo signalling (Extended Data Fig. 10a–g). In addition, in epithelial cancer cells, decreased levels of ECAD, NF2 or Hippo pathway activity, and enhanced activation of YAP can promote epithelial–mesenchymal transition (EMT) and metastasis[19]. Consistently, as TGFβ can induce the expression of several EMT genes, it also enhanced ferroptosis in mammary tumour cells isolated from MMTV-neu mice at high cell density (Extended Data Fig. 10h–j).

Collectively, we describe a non-cell-autonomous mechanism for the regulation of ferroptosis: neighbouring cells can have a considerable effect on the decision-making of ferroptosis via the

cadherin–NF2–Hippo–YAP signalling axis. Considering that multicellular organisms are under frequent insult of oxidative stress, this intercellular mechanism might represent another layer of crucial defence to protect themselves from ferroptosis, a terminal consequence of oxidative stress.

Because cellular metabolism has a crucial role in ferroptosis, and enhanced proliferation often leads to stronger metabolism, it is possible that proliferation-stimulating oncogenic mutation may be a good predictor of ferroptosis sensitivity. However, previous publications argue against this view. For example, loss of function of the tumour suppressors p53 and BAP1 increases resistance, instead of sensitivity, to ferroptosis[5,7]. Furthermore, unlike YAP(S127A), overexpression of the oncogenic PIK3CA(H1047R) mutant did not sensitize confluent 211H cells to ferroptosis, although both increased proliferation (Extended Data Fig. 10k–m). Together, oncogenic mutations may affect ferroptosis by mechanisms other than enhancing proliferation.

As the cadherin–NF2–Hippo–YAP signalling axis is frequently mutated in cancer, this study has clear implications for cancer therapies—malignant alterations of several components in this signalling axis all sensitize cancer cells to ferroptosis. A potential concern about the feasibility of ferroptosis-inducing cancer therapy is whether there is any selectivity of the ferroptosis-inducing agents towards cancer cells compared with normal tissue. Our finding suggests that there might be a dose-responsive window for cancers that contain certain genetic signatures and that ferroptosis-inducing cancer therapies—if available (IKE and sorafenib hold potential for this purpose)—might have considerable benefits in overcoming cancer resistance to current treatments.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41586-019-1426-6.

Received: 7 November 2018; Accepted: 27 June 2019;
Published online 24 July 2019.

1. Stockwell, B. R. et al. Ferroptosis: a regulated cell death nexus linking metabolism, redox biology, and disease. *Cell* **171**, 273–285 (2017).
2. Gao, M. & Jiang, X. To eat or not to eat—the metabolic flavor of ferroptosis. *Curr. Opin. Cell Biol.* **51**, 58–64 (2018).
3. Yang, W. S. et al. Regulation of ferroptotic cancer cell death by GPX4. *Cell* **156**, 317–331 (2014).
4. Gao, M. et al. Role of mitochondria in ferroptosis. *Mol. Cell* **73**, 354–363.e3 (2019).
5. Jiang, L. et al. Ferroptosis as a p53-mediated activity during tumour suppression. *Nature* **520**, 57–62 (2015).
6. Jennis, M. et al. An African-specific polymorphism in the *TP53* gene impairs p53 tumor suppressor function in a mouse model. *Genes Dev.* **30**, 918–930 (2016).
7. Zhang, Y. et al. BAP1 links metabolic regulation of ferroptosis to tumour suppression. *Nat. Cell Biol.* **20**, 1181–1192 (2018).
8. Viswanathan, V. S. et al. Dependency of a therapy-resistant state of cancer cells on a lipid peroxidase pathway. *Nature* **547**, 453–457 (2017).
9. Hangauer, M. J. et al. Drug-tolerant persister cancer cells are vulnerable to GPX4 inhibition. *Nature* **551**, 247–250 (2017).
10. Hmeljak, J. et al. Integrative molecular characterization of malignant pleural mesothelioma. *Cancer Discov.* **8**, 1548–1565 (2018).
11. Bueno, R. et al. Comprehensive genomic analysis of malignant pleural mesothelioma identifies recurrent mutations, gene fusions and splicing alterations. *Nat. Genet.* **48**, 407–416 (2016).
12. Seiler, A. et al. Glutathione peroxidase 4 senses and translates oxidative stress into 12/15-lipoxygenase dependent- and AIF-mediated cell death. *Cell Metab.* **8**, 237–248 (2008).
13. Schneider, M. et al. Absence of glutathione peroxidase 4 affects tumor angiogenesis through increased 12/15-lipoxygenase activity. *Neoplasia* **12**, 254–263 (2010).
14. Gao, M., Monian, P., Quadri, N., Ramasamy, R. & Jiang, X. Glutaminolysis and transferrin regulate ferroptosis. *Mol. Cell* **59**, 298–308 (2015).
15. van Roy, F. & Berx, G. The cell–cell adhesion molecule E-cadherin. *Cell. Mol. Life Sci.* **65**, 3756–3788 (2008).
16. Kim, N. G., Koh, E., Chen, X. & Gumbiner, B. M. E-cadherin mediates contact inhibition of proliferation through Hippo signaling-pathway components. *Proc. Natl Acad. Sci. USA* **108**, 11930–11935 (2011).
17. Okada, T., Lopez-Lago, M. & Giancotti, F. G. Merlin/NF-2 mediates contact inhibition of growth by suppressing recruitment of Rac to the plasma membrane. *J. Cell Biol.* **171**, 361–371 (2005).
18. Zhao, B., Lei, Q. Y. & Guan, K. L. The Hippo–YAP pathway: new connections between regulation of n size and cancer. *Curr. Opin. Cell Biol.* **20**, 638–646 (2008).
19. Pan, D. The Hippo signaling pathway in development and cancer. *Dev. Cell* **19**, 491–505 (2010).
20. Li, W., Cooper, J., Karajannis, M. A. & Giancotti, F. G. Merlin: a tumour suppressor with functions at the cell cortex and in the nucleus. *EMBO Rep.* **13**, 204–215 (2012).
21. Li, W. et al. Merlin/NF2 loss-driven tumorigenesis linked to CRL4(DCAF1)-mediated inhibition of the hippo pathway kinases Lats1 and 2 in the nucleus. *Cancer Cell* **26**, 48–60 (2014).
22. Dupont, S. et al. Role of YAP/TAZ in mechanotransduction. *Nature* **474**, 179–183 (2011).
23. Falk, M. H. et al. Apoptosis in Burkitt lymphoma cells is prevented by promotion of cysteine uptake. *Int. J. Cancer* **75**, 620–625 (1998).
24. Varelas, X. & Wrana, J. L. Coordinating developmental signaling: novel roles for the Hippo pathway. *Trends Cell Biol.* **22**, 88–96 (2012).
25. Zhao, B. et al. Inactivation of YAP oncoprotein by the Hippo pathway is involved in cell contact inhibition and tissue growth control. *Genes Dev.* **21**, 2747–2761 (2007).
26. Doll, S. et al. ACSL4 dictates ferroptosis sensitivity by shaping cellular lipid composition. *Nat. Chem. Biol.* **13**, 91–98 (2017).
27. Zhang, Y. et al. Imidazole ketone erastin induces ferroptosis and slows tumor growth in a mouse lymphoma model. *Cell Chem. Biol.* **26**, 623–633.e9 (2019).
28. Dixon, S. J. et al. Pharmacological inhibition of cystine–glutamate exchange induces endoplasmic reticulum stress and ferroptosis. *eLife* **3**, e02523 (2014).
29. Dubey, S. et al. A phase II study of sorafenib in malignant mesothelioma: results of Cancer and Leukemia Group B 30307. *J. Thorac. Oncol.* **5**, 1655–1661 (2010).
30. Papa, S. et al. Phase 2 study of sorafenib in malignant mesothelioma previously treated with platinum-containing chemotherapy. *J. Thorac. Oncol.* **8**, 783–787 (2013).

METHODS

Cell culture. MEFs, mouse NF639 cells, human epithelial tumour cells, and human mesothelioma cells were cultured in DMEM containing 10% fetal calf serum (FCS), 2 mM L-glutamine, 100 U ml^{-1} penicillin and 100 $\mu g\ ml^{-1}$ streptomycin. CA-46 Burkitt lymphoma cells were cultured in RPMI medium supplemented with 20% serum and 100 U ml^{-1} penicillin and 100 $\mu g\ ml^{-1}$ streptomycin. The mesothelioma cell line panel was a gift from the Giancotti Laboratory. Media were prepared by the MSKCC Media Preparation Core Facility. All cell lines were subjected to STR authentication through ATCC and were tested for mycoplasma contamination.

Generation of 3D spheroids. Spheroids were generated by plating tumour cells at 10^3 per well into U-bottom Ultra Low Adherence (ULA) 96-well plates (Corning). Optimal 3D structures were achieved by centrifugation at 600*g* for 5 min followed by addition of 2.5% (v/v) Matrigel (Corning). Plates were incubated for 72 h at 37 °C, 5% CO_2, 95% humidity for formation of a single spheroid of cells. Spheroids were then treated with erastin in fresh medium containing Matrigel for the indicated time.

Induction and inhibition of ferroptosis. To induce ferroptosis, cells with different density were seeded in 6-well plates. For cystine-starvation experiments, cells were washed twice with PBS and then cultured in cystine-free medium in the presence of 10% (v/v) dialysed FBS for the indicated time. The ferroptosis-inducing compounds erastin and RSL3 and the ferroptosis inhibitor ferrostatin-1 were purchased from Sigma-Aldrich.

Measurement of cell death, cell viability and lipid peroxidation. Cell death was analysed by staining for propidium iodide (Invitrogen) or SYTOX Green (Invitrogen) followed by microscopy or flow cytometry. For 3D spheroids, cell viability was determined using the CellTiter-Glo 3D Cell Viability Assay (Promega) according to the manufacturer's instructions. Viability was calculated by normalizing ATP levels to spheroids treated with normal medium. To analyse lipid peroxidation, cells were stained 5 μM BODIPY-C11 (Invitrogen) for 30 min at 37 °C followed by flow cytometric analysis. Lipid ROS-positive cells are defined as cells with FITC fluorescence greater than 99% of the unstained sample.

Immunoblotting. Nuclear and non-nuclear (membranes and cytosol) fractions were prepared as previously described. Proteins in the cell lysate were resolved on 8% or 15% SDS–PAGE gels and transferred to a nitrocellulose membrane. Membranes were incubated in 5% skim milk for 1 h at room temperature and then with primary antibodies diluted in blocking buffer at 4 °C overnight. The following primary antibodies were used: rabbit anti-GPX4, mouse anti E-cadherin, rabbit anti-N-cadherin, rabbit anti-NF2/Merlin, rabbit anti-transferrin receptor (Abcam, Cambridge), mouse anti-β-actin, mouse anti-Flag, mouse anti-HA (Sigma-Aldrich), rabbit anti-NF2/Merlin, rabbit anti-phospho-NF2/Merlin (Ser518), rabbit anti-LATS1, rabbit anti-LATS2, rabbit anti-YAP, rabbit anti-phospho-YAP (Ser127), mouse anti-CAS9, rabbit anti-p110α, mouse anti-AKT, rabbit anti-phospho-AKT (Ser473), rabbit anti-TAZ, rabbit anti-pan cadherin (Cell Signaling), rabbit anti-ACSL4 (Thermo Fisher), mouse anti-α-tubulin (Calbiochem), rabbit anti-GFP (Invitrogen). Goat anti-mouse or donkey anti-rabbit IgG (Invitrogen) conjugated to horseradish peroxidase (HRP) and an Amersham Imager 600 (GE Healthcare Life Sciences) was used for detection. Representative blots of at least two independent experiments are shown. After three washes, the membranes were incubated with goat anti-mouse HRP-conjugated antibody or donkey anti-rabbit HRP-conjugated antibody at room temperature for 1 h and subjected to chemiluminescence using Clarity Western ECL Substrate (Bio-Rad).

Plasmids and cloning. pWZL Blast mouse E-cadherin and pWZL Blast DN E-cadherin were from the Weinberg Laboratory (Addgene plasmids 18804 and 18800, respectively). pRK5-Flag-HA-NF2 was from the Giancotti laboratory (Addgene plasmid 27104). The 8xGTIIC-luciferase reporter was from the Piccolo laboratory (Addgene plasmid 34615). mCherry-TFR-20 was from the Davidson laboratory (Addgene plasmid 55144). pQCXIH-Flag-YAP-S127A was from the Guan laboratory (Addgene plasmid 33092). pBABE-Flag-HA-NF2 was generated by PCR from pRK5-Flag-HA-NF2 (primers listed in Supplementary Table 2), digested by PacI and EcoRI FastDigest restriction enzymes (Thermo Fisher), and ligated into the empty pBABE-puro backbone using T4 ligase (NEB). FUW-tetO-Flag-HA-NF2 was created by digesting pRK5-Flag-HA-NF2 with EcoRI and XbaI and was ligated into the FUW-tetO-MCS vector from the Piccolo laboratory (Addgene plasmid 84008). FUW-m2rtTA was from the Jaenisch laboratory (Addgene plasmid 20342). PIK3CA(H1047R) was a gift from the Cantley laboratory (Weill Cornell Medicine).

Gene silencing and expression. Lentiviral vectors encoding shRNAs targeting human *ECAD*, human *NCAD*, human and mouse *Nf2*, human *LATS1* and *LATS2*, and human *TFRC* were generated by the core facility of MSKCC and are listed in Supplementary Table 1. Lentiviruses were produced by the co-transfection of the lentiviral vector with the Delta-VPR envelope and CMV VSV-G packaging plasmids into 293T cells using PEI. Medium was changed 12 h after transfection. The supernatant was collected 48 h after transfection and passed through a 0.45-μm filter to eliminate cells. Cells were incubated with infectious particles in the presence of 4 $\mu g\ ml^{-1}$ polybrene (Sigma-Aldrich) overnight and cells were given fresh complete medium. After 48 h, cells were placed under the appropriate antibiotic selection.

Generation of constitutive and inducible CRISPR–Cas9-mediated gene knockouts. ECAD, YAP- and ACSL4-depleted cells were generated using the CRISPR–Cas9-mediated knockout system. HCT116 cells were transfected with a human *ECAD* CRISPR–Cas9 knockout plasmid (sc-400031), and HCT116-shNF2 cells were transfected with a human *YAP* CRISPR–Cas9 knockout plasmid (sc-400040) or a human *ACSL4* CRISPR–Cas9 KO plasmid (sc-401649), all purchased from Santa Cruz Biotechnology. The target sequence was a pool of three different gRNA plasmids located within the coding DNA sequence fused to *Streptococcus pyogenes Cas9* and GFP. Single GFP^+ cells were sorted using a BD FACSAria II cytometer (BD Biosciences) and a 96-well plate, and single-cell clones were tested by western blotting.

The lentiviral Dox-inducible Flag–Cas9 vector pCW-Cas9 and pLX-sgRNA were from E. Lander and D. Sabatini (Addgene plasmids 50661 and 50662, respectively). Guide RNA sequence CACGCCCGATACGCTGAGTG was used to target human *GPX4*. To construct the lentiviral sgRNA vector for *GPX4*, a pair of oligonucleotides (forward and reverse) was annealed, phosphorylated and ligated into pLX-sgRNA. Lentiviral particles containing the sgRNA or Cas9 vectors were produced by co-transfection of the vectors with the Delta-VPR envelope and CMV VSV-G packaging plasmids into 293T cells using PEI. Medium was changed 12 h after transfection and supernatant was collected 48 h after transfection. MSTO-211H cells in 6-well tissue culture plates were infected in pCW-Cas9 virus-containing supernatant containing 4 $\mu g\ ml^{-1}$ of polybrene. Twenty-four hours after infection, virus was removed, and cells were selected with 2 $\mu g\ ml^{-1}$ puromycin. Single clones were screened for inducible Cas9 expression. Dox (2 $\mu g\ ml^{-1}$) was added to the culture medium for 3 days. Single clones with Cas9 expression were infected with *GPX4* gRNA virus-containing supernatant containing 8 $\mu g\ ml^{-1}$ polybrene. Twenty-four hours after infection, virus was removed and cells were selected with 10 $\mu g\ ml^{-1}$ blasticidin. Single clones with Dox-inducible Cas9 expression and *GPX4* knockout were amplified for further experiments, named GPX4-iKO MSTO-211H cells.

ChIP assay. Cells were crosslinked in 0.75% formaldehyde for 15 min, and glycine was added to a final concentration of 125 mM for 5 min. After washing with cold PBS, cells were collected in PBS and sonicated on an ultrasonic homogenizer for 10 min at 20% power on ice to shear DNA to an average fragment size of 200–1,000 bp. Fifty microlitres of each sonicated sample was removed to determine the DNA concentration and fragment size. Cell lysates were incubated overnight with 20 μl Magna ChIP Protein A+G Magnetic Beads (EMD Millipore) and 10 μg ChIP grade TEAD4 antibody (Abcam) at 4 °C. Beads were collected, washed and treated with proteinase K for 2 h at 60 °C and RNase for 1 h at 37 °C. DNA was purified with a PCR purification kit (Qiagen). DNA fragments were assessed by quantitative PCR with reverse transcription (qRT–PCR) using the primer sequences listed in Supplementary Table 2. Samples were normalized to input DNA.

RNA extraction and qRT–PCR. RNA was extracted using the TRIzol reagent (Invitrogen). Samples were treated with chloroform (20%), vortexed briefly, and incubated at room temperature for 15 min. Samples were then centrifuged at high speed at 4 °C for 15 min. The aqueous phase was moved to a new tube and an equal volume of isopropanol was added. Samples were incubated at room temperature for 10 min, followed by centrifugation at high speed at 4 °C for 10 min. Pellets were washed in 95% ethanol, dried and resuspended in nuclease-free water. cDNA was synthesized using iScript cDNA Synthesis Kit according to the manufacturer's instructions (Bio-Rad). qRT–PCR was performed with IQ SYBR Green Supermix (Bio-Rad) in a CFX Connect Real-Time PCR Detection System (Bio-Rad). Primer sequences are listed in Supplementary Table 2.

In vivo xenograft mouse study. GPX4-iKO MSTO-211H cells were infected with lentiviral vectors encoding shRNAs targeting human *NF2* or non-targeting control (shNT) (GeneCopoeia). The resulting cells were called 'shNT-GPX4-iKO' and 'shNF2-GPX4-iKO' MSTO-211H cells. Six- to eight-week-old female athymic *nu/nu* mice were purchased from Envigo. For subcutaneous tumour models, mice were injected in the right flank with 1×10^7 shNT-GPX4-iKO or shNF2-GPX4-iKO MSTO-211H cells suspended in 150 μl Matrigel. Tumours were measured with callipers every three days. When tumours reached a mean volume of 100 mm^3, mice with similarly sized tumours were grouped into four treatment groups. For control or knockout cohorts, mice were given intraperitoneal injections of 0.9% sterile saline or Dox (100 mg kg^{-1} body weight) for two days. At the same time, mice were provided with either a normal or a Dox diet for control or knockout cohorts, respectively. For all experiments, mice were killed at a pre-determined endpoint. According to the Institutional Animal Care and Use Committee (IACUC) protocol for these experiments, once any tumour exceeded a volume of 1,000 mm^3, 1.5 cm in diameter or 10% of body weight, the mice would immediately be euthanized. At the end of the study, mice were euthanized with CO_2 and tumours were taken for immunohistochemical staining. Results are presented as mean tumour volume ± s.d.

For shLATS1/2 subcutaneous tumour models, female athymic *nu/nu* mice aged 6–8 weeks were injected in the right flank with 2×10^6 shNT HCT116 cells or shLATS1/2 HCT116 cells. Tumours were measured with callipers daily. When tumours reached a mean volume of 90 mm^3, mice were randomized into 4 groups and treated with vehicle (65% D5W (5% dextrose in water), 5% Tween-80, 30% PEG-400) or 50 mg kg^{-1} IKE (65% D5W (5% dextrose in water), 5% Tween-80, 30% PEG-400) via intraperitoneal injection once a day. At the end of the study, mice were euthanized with CO_2 and tumours were taken for measurement of weight. According to the IACUC protocol for these experiments, once any tumour exceeded a volume of 1,000 mm^3, 1.5 cm in diameter or 10% of body weight, the mice would immediately be euthanized.

All protocols for mouse experiments were approved by the Memorial Sloan Kettering Cancer Center IACUC.

Orthotopic pleural mesothelioma mouse model. shNT-GPX4-iKO and shNF2-GPX4-iKO MSTO-211H cells were infected with retroviral TK-GFP-luciferase (TGL) reporter vector. To develop the orthotopic mouse model of pleural mesothelioma, female NOD/SCID mice (Envigo) aged 6–8 weeks were used. Mice were anaesthetized using inhaled isoflurane and oxygen. Intrapleural injection of 2×10^6 shNT-GPX4-iKO-TGL or shNF2-GPX4-iKO-TGL MSTO-211H cells in 100 μl of serum-free medium via a left thoracic incision was performed to establish the orthotopic mesothelioma tumour model. Tumour growth was monitored by weekly BLI for luciferase and mice were monitored daily for survival. NOD/SCID mice bearing tumours were anaesthetized using isoflurane and injected intraperitoneally with 50 mg kg^{-1} D-luciferin (Molecular Probes). BLI was measured with 18 filters (500–840 nm) in an IVIS Spectrum (PerkinElmer) 10 min after injection. During image acquisition, mice were maintained on isoflurane via nose cone. Bioluminescence images were acquired using an IVIS Spectrum. The BLI signal was reported as total flux (photons per second), which represents the average of ventral and dorsal flux. At the end-point of the study, the mice were injected with D-luciferase and euthanized 10 min later. Organs were exposed and the BLI signal was measured. After organs were excised, BLI images were taken again as described. Imaging analysis was performed using the Living Image software (Caliper Life Sciences) All protocols for mouse experiments were approved by the Memorial Sloan Kettering Cancer Center IACUC.

Immunohistochemistry. Formalin-fixed, paraffin-embedded specimens were collected, and a routine H&E slide was first evaluated. Immunohistochemical staining was done on 5-μm-thick paraffin-embedded sections using mouse anti-NF2/Merlin (Abcam), rabbit anti-GPX4 (Abcam), rabbit anti-PTGS2 (Cell Signaling), mouse anti-Ki67 (Cell Signaling), rabbit anti-ACSL4 (Thermo Fisher), rabbit anti-TFRC (Abcam) and rabbit anti-YAP (Cell Signaling) antibodies with a standard avidin-biotin HRP detection system according to manufacturer's instructions (anti-mouse/rabbit HRP-DAB Cell & Tissue Staining Kit, R&D Systems). Tissues were counterstained with haematoxylin, dehydrated and mounted. In all cases, antigen retrieval was done with the BD Retrievagen Antigen Retrieval Systems as per manufacturer's instructions.

Tumour spheroid invasion assay. Spheroids were generated as described in 200 μl complete growth medium and cultured for 72 h after cell seeding. The ULA 96-well plates containing 3-day-old spheroids were placed on ice. One-hundred microlitres per well of growth medium was removed from the spheroid plates. Using ice-cold tips, 100 μl of Matrigel was transferred to each well and mixed gently with medium, avoiding disturbance of the spheroids. Plates were placed in an incubator at 37 °C to allow the Matrigel to solidify. One hour later, 100 μl per well of complete growth medium was added. Images for each tumour spheroid were taken 48 h later.

Statistical analysis. All statistical analyses were performed using GraphPad Prism 6.0 Software. Data are presented as mean ± s.d. from three independent experiments. *P* values were determined by Student's two-tailed *t*-test, one-way ANOVA or two-way ANOVA as indicated in the figure legends. For ANOVA, adjustments were made for multiple comparisons by Dunnett or Tukey corrections as appropriate. Exact *P* values can be found in figure legends. In cases where more than one comparison has the same statistical range, values are listed as they appear from left to right in the corresponding panel. No statistical methods were used to predetermine sample size. Unless stated otherwise, the experiments were not randomized and investigators were not blinded to allocation during experiments and outcome assessment.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

For western blot source data, see Supplementary Fig. 1. For the gating strategy used for flow cytometry experiments, see Supplementary Fig. 2. Raw data for all experiments are available as Source Data to the relevant figures. ChIP–seq datasets analysed in this article are publicly available in the ENCODE database under the identifiers GSM1010875 and GSM1010868.

Acknowledgements We thank E. De Stanchina and E. Peguero for their help with mouse modelling experiments. We thank members of the Jiang laboratory for critical reading and suggestions. This work is supported by the National Institutes of Health (NIH) R01CA204232 (to X.J.), a Geoffrey Beene Cancer Research fund (to X.J.), a Functional Genomic Initiative fund (to X.J.), a China Scholarship Council fellowship (to J.W.), and NIH T32 fellowship 5T32GM008539-23 (to A.M.M.), National Cancer Institute R35CA209896 and P01CA087497 (to B.R.S.), National Natural Science Foundation of China 31871388 (to M.G.). This work is also supported by NCI cancer centre core grant P30 CA008748 to Memorial Sloan Kettering Cancer Center.

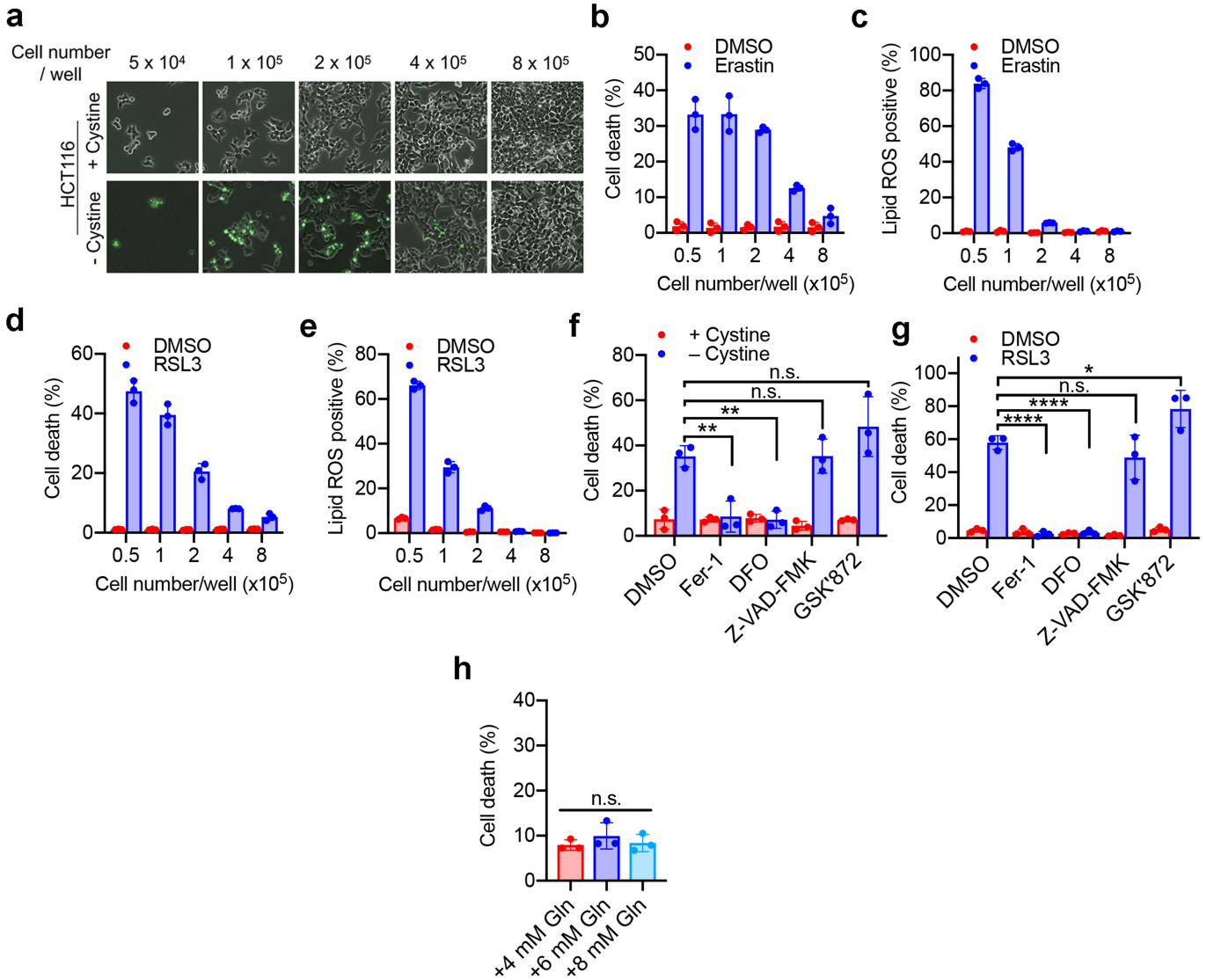
Author contributions J.W., A.M.M. and X.J. conceived the original idea and designed the study. J.W. and A.M.M. performed most experiments. M.G., H.B. and Y.L. generated several reagents and the inducible GPX4 knockout (GPX4-iKO) used in mouse experiments. B.R.S. supplied IKE and protocols for IKE administration to mice. Z.-N.C. and X.J. supervised the research. J.W., A.M.M., Z.-N.C. and X.J. wrote the paper.

Competing interests B.R.S. holds equity in and serves as a consultant to Inzen Therapeutics, consults with GLG and Guidepoint Global, and is an inventor on patents and patent applications related to IKE and ferroptosis.

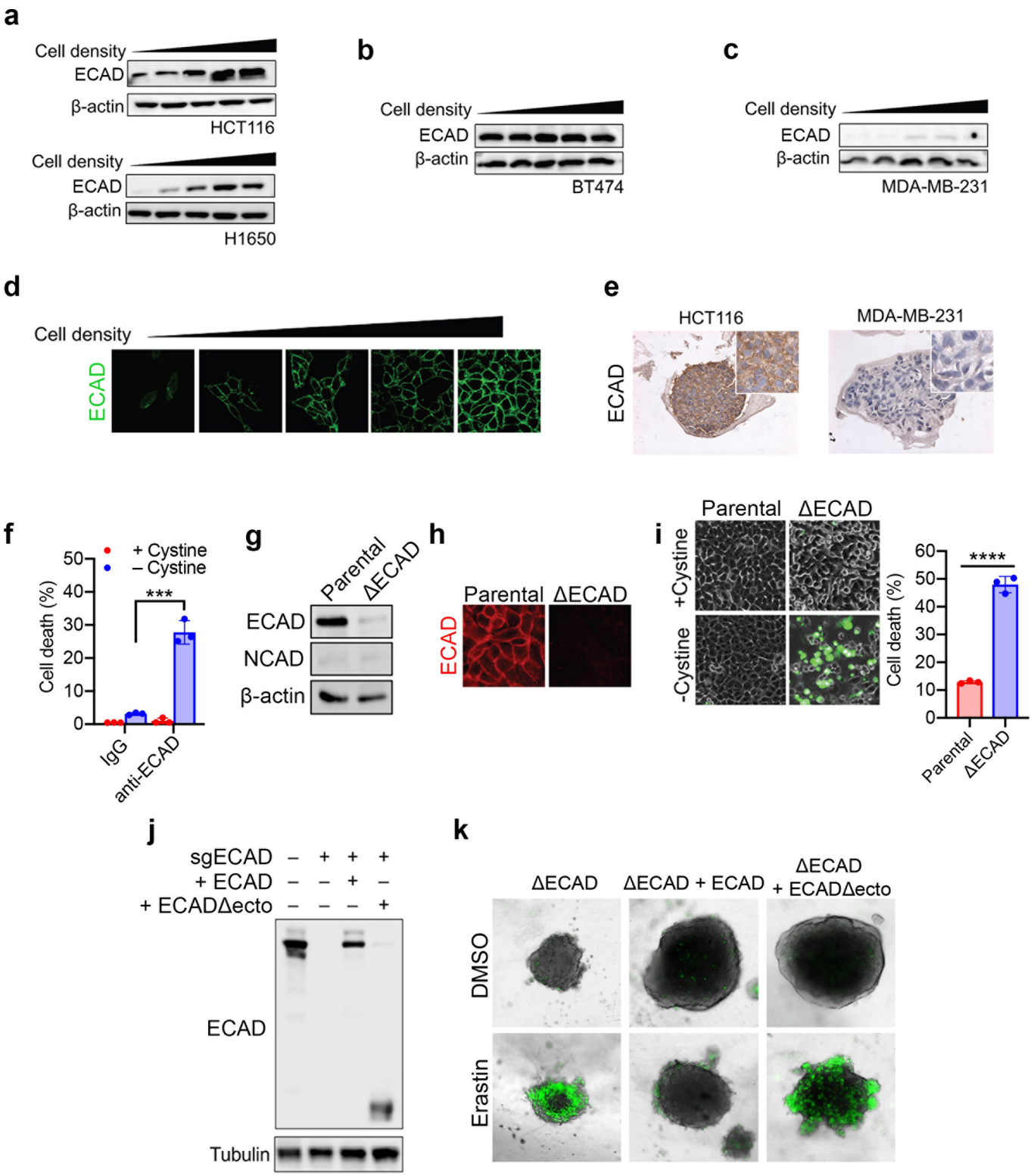
Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1426-6.

Correspondence and requests for materials should be addressed to Z.-N.C. or X.J.

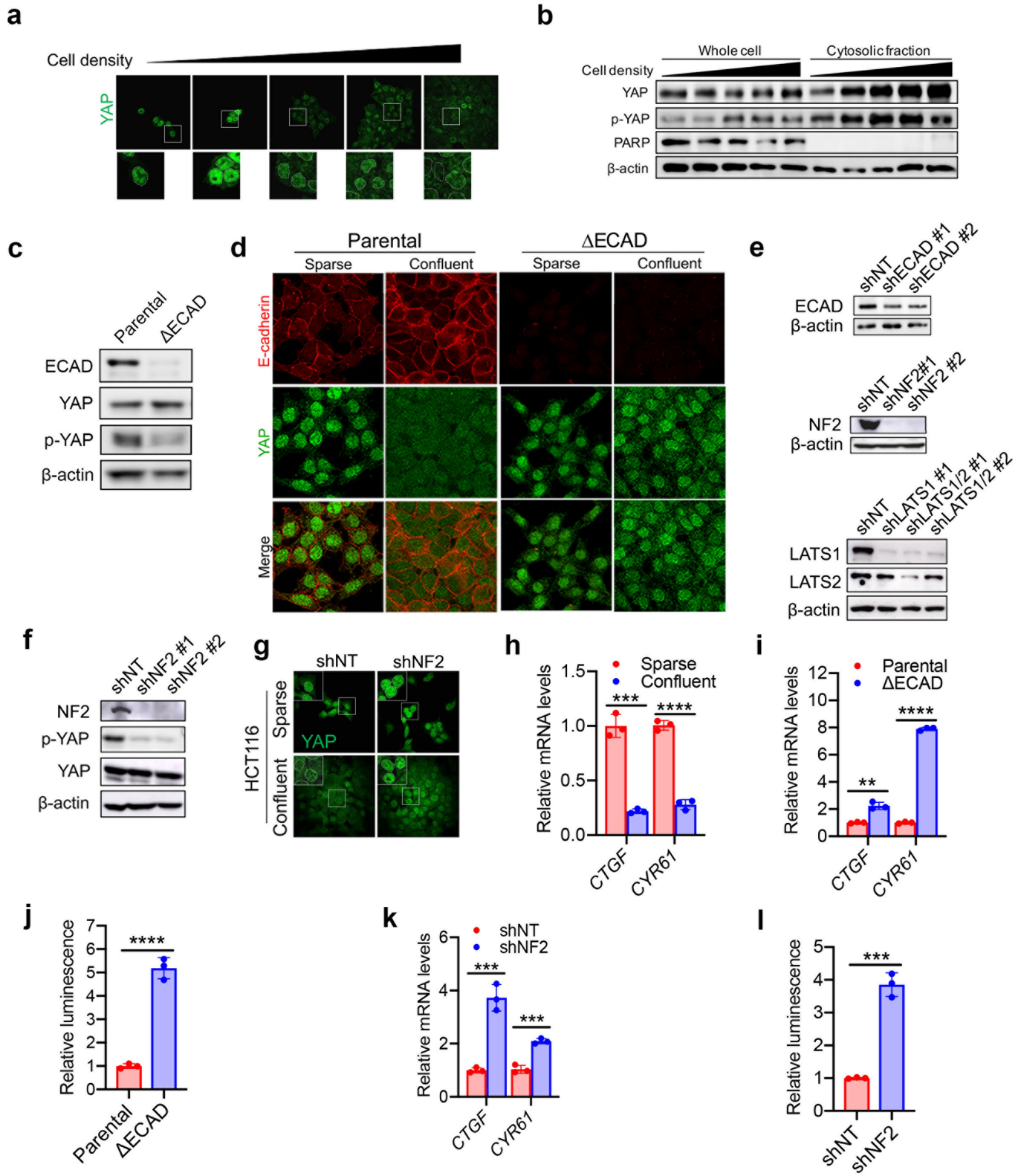


Extended Data Fig. 1 | Intercellular contact suppresses ferroptosis. **a**, **b**, HCT116 cells were seeded at the indicated density in 6-well plates and cultured for 24 h. **a**, Ferroptosis was measured by SYTOX Green staining after cystine starvation for 30 h. Phase contrast and fluorescent images are overlaid (original magnification, ×100). **b**, Cell death was measured in HCT116 cells at different densities treated with 30 μM erastin for 30 h, quantified by SYTOX Green staining followed by flow cytometry. **c**, Lipid ROS production of cells in **b** was assessed by C11-BODIPY staining followed by flow cytometry after 24 h of erastin treatment **d**, Cell death was measured in HCT116 cells cultured at the indicated cell densities and treated with 5 μM RSL3 for 24 h. **e**, Lipid ROS production in HCT116 cells cultured at the indicated cell densities and treated with 5 μM RSL3 for 16 h. **f**, HCT116 cells were seeded at 5×10^4 cells per well, grown for 24 h, and treated with: 1 μM Fer-1; 50 μg ml^{-1} of the iron chelator deferoxamine (DFO); 20 μM of the pan-caspase inhibitor Z-VAD-FMK; or 10 μM of the RIPK3 inhibitor GSK'872, in complete medium or cysteine-free medium for 30 h, followed by cell death measurement. n.s., $P = 0.9999$, 0.1995 (left to right). $^{**}P = 0.0070$, 0.0050; one-way ANOVA. **g**, Cell death was measured in HCT116 cells seeded at 5×10^4 cells per well, grown for 24 h and treated with 5 μM RSL3 or DMSO and inhibitors as in **f** for 24 h. n.s., $P = 0.4989$. $^{*}P = 0.0366$, $^{****}P < 0.0001$; one-way ANOVA. **h**, Cell death analysis in HCT116 cells seeded at 8×10^5 cells per well, grown for 24 h and treated with cystine-free medium containing the indicated amounts of glutamine for 30 h. Cell death was measured by SYTOX Green staining followed by flow cytometry. n.s., $P = 0.5156$; one-way ANOVA. All data are mean ± s.d. from $n = 3$ biological replicates.



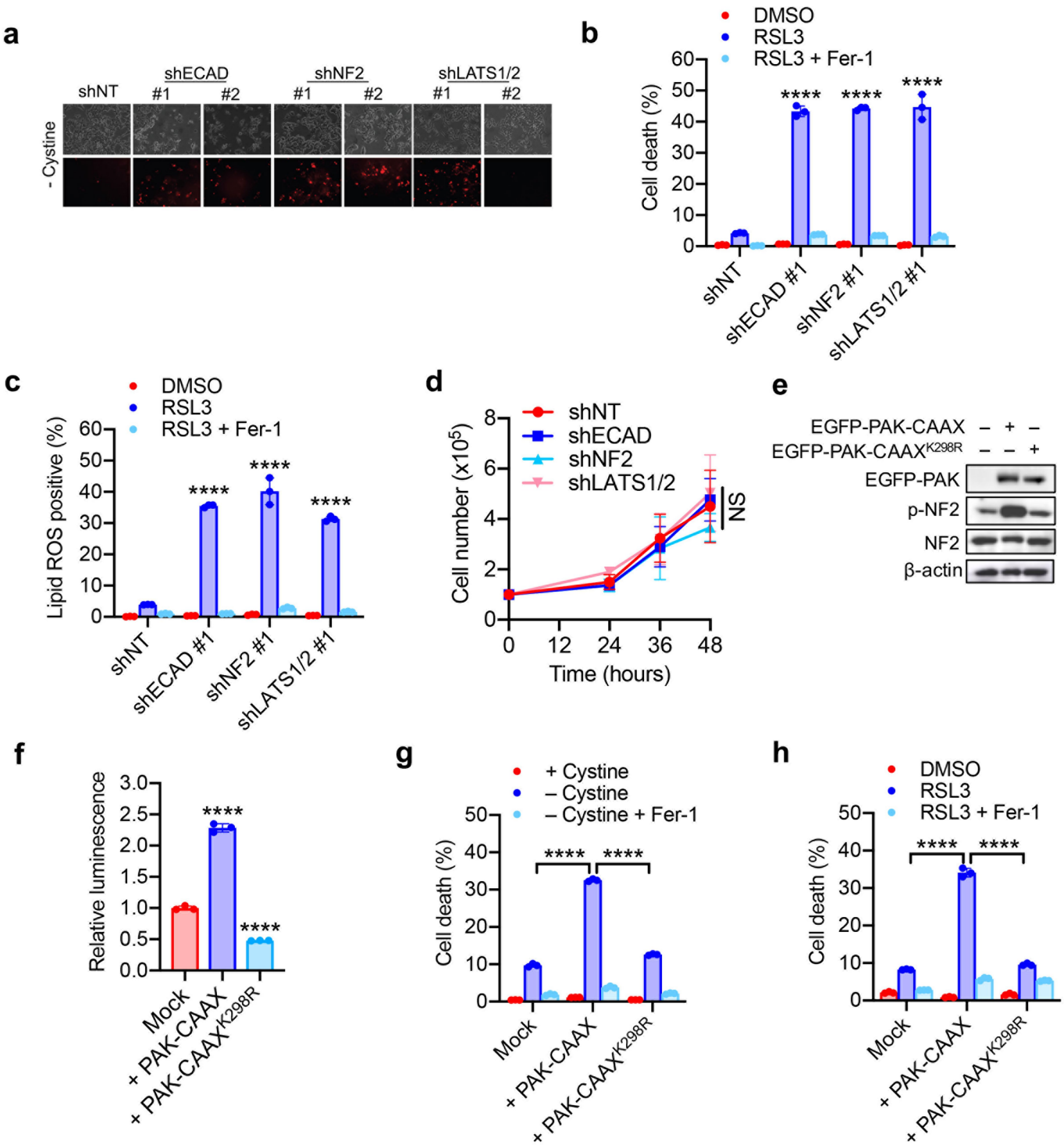
Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | ECAD-mediated intercellular interaction regulates ferroptosis in a density-dependent manner. a, ECAD expression increases with cell density in HCT116 (top) and H1650 (bottom) cells. **b**, BT474 cells express high levels of ECAD regardless of cell density. **c**, MDA231 cells express low levels of ECAD regardless of cell density. **d**, Immunofluorescence of ECAD at increasing cell density in HCT116 cells (original magnification, ×200). **e**, Tumour spheroids generated from HCT116 or MDA231 cells were fixed, sectioned and stained for ECAD expression by immunohistochemistry (original magnification, ×100). **f**, HCT116 cells were treated with either IgG or an anti-ECAD antibody that blocks dimerization. Cell death was measured by propidium iodide staining followed by flow cytometry after cystine starvation for 30 h. ***$P = 0.0003$; two-tailed *t*-test. **g**, Western blot analysis of the expression of ECAD and NCAD in HCT116 cells after CRISPR–Cas9-mediated ECAD depletion (ΔECAD). **h**, ECAD depletion in HCT116 cells was further confirmed by immunofluorescence (magnification 200×). **i**, Cell death measurement of ΔECAD and the parental cell line seeded at a density of 4×10^5 cells per well after cystine starvation for 30 h. Original magnification, ×100. ****$P < 0.0001$; two tailed *t*-test. **j**, Western blot analysis confirming reconstitution of ECAD or ectodomain-truncated ECAD (ECADΔecto) into ECAD-depleted HCT116 cells. **k**, ΔECAD cells or ΔECAD cells re-expressing full-length ECAD or ECADΔecto were cultured into spheroids and treated with erastin for 30 h, followed by SYTOX Green staining to monitor cell death (magnification 100×). All data are mean ± s.d. from $n = 3$ biological replicates.

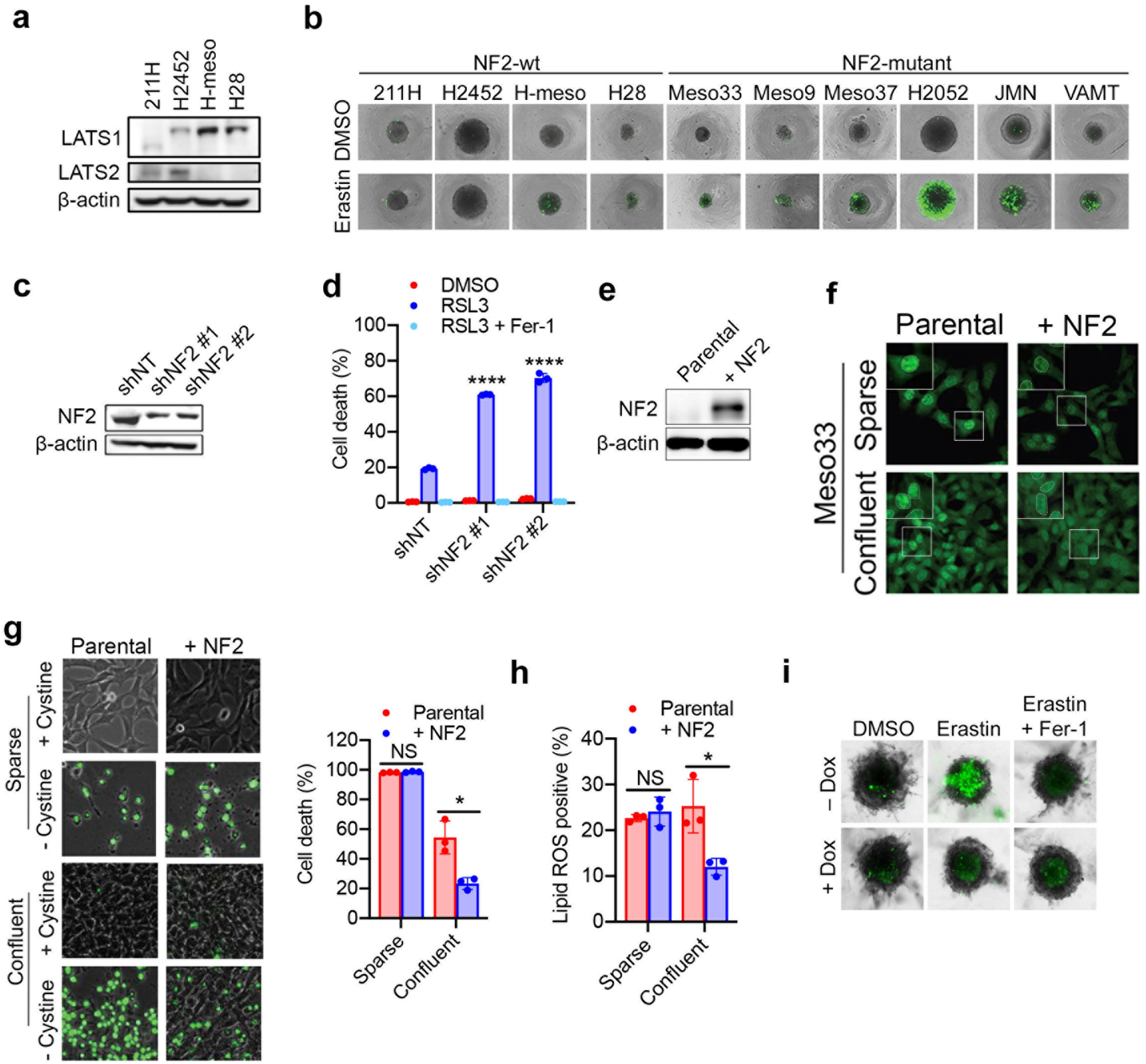


Extended Data Fig. 3 | See next page for caption.

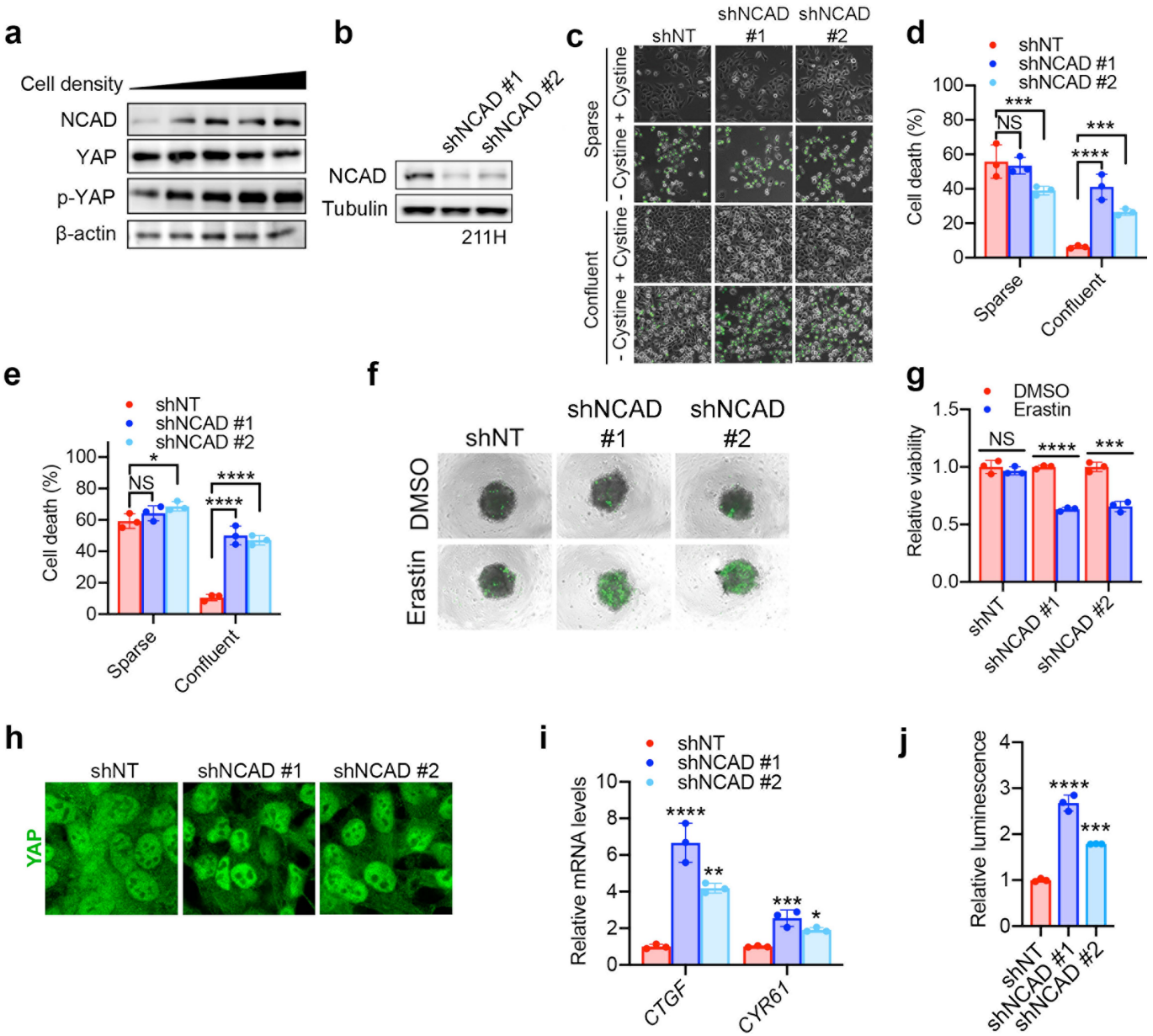
Extended Data Fig. 3 | Cell density, ECAD and NF2 converge on the transcriptional co-regulator YAP. a, HCT116 cells were cultured at different cell densities and YAP localization was assessed by immunofluorescence. Original magnification, ×200. Bottom images are enlarged to show localization. **b**, Western blot analysis of phosphorylated YAP (p-YAP; at Ser127) and YAP in whole-cell or cytosolic fractions of HCT116 cells cultured at different densities. PARP was used as a marker for nuclear protein. **c**, Western blot analysis of ECAD, YAP and p-YAP in parental and ΔECAD HCT116 cells. **d**, Immunofluorescence of YAP (green) and ECAD (red) in parental and ΔECAD HCT116 cells. Original magnification, ×400. **e**, Western blot analysis confirming knockdown efficiency of ECAD (shECAD #1 and #2), NF2 (shNF2 #1 and #2), or LATS1 or LATS2 (shLATS1/2 #1 and #2, and shLATSs1 #1) in HCT116 cells. **f**, Western blot analysis of NF2, p-YAP and YAP in HCT116 cells transfected with shNT and shNF2. **g**, Knockdown of NF2 in HCT116 cells induced the nuclear accumulation of YAP in dense cell cultures, as assessed by immunofluorescence. Original magnification, ×200. **h**, Transcriptional levels of the canonical YAP targets *CTGF* and *CYR61* by qPCR in HCT116 cells seeded at 1×10^5 (sparse) or 8×10^5 (confluent) cells per well in 6-well dishes and grown for 24 h. $^{***}P = 0.0002$, $^{****}P < 0.0001$; two-tailed *t*-test. **i**, Transcription levels of *CTGF* and *CYR61* measured by qPCR in parental and ΔECAD HCT116 cells plated at high density. $^{**}P = 0.0013$, $^{****}P < 0.0001$; two-tailed *t*-test. **j**, YAP/TEAD transcriptional activity in HCT116 and ΔECAD cells was measured by a luciferase assay using the 8xGTIIC-luciferase reporter. $^{****}P < 0.0001$; two-tailed *t*-test. **k**, Transcription levels of *CTGF* and *CYR61* measured by qPCR in HCT116 cells plated at high density and transfected with shNT or shNF2. $^{***}P = 0.0007$, 0.0005 (left to right); two-tailed *t*-test. **l**, YAP/TEAD transcriptional activity in HCT116 cells transfected with shNT and shNF2 cells was measured by a luciferase assay using the 8xGTIIC-luciferase reporter. $^{***}P = 0.0002$; two-tailed *t*-test. All data are mean ± s.d. from $n = 3$ biological replicates.



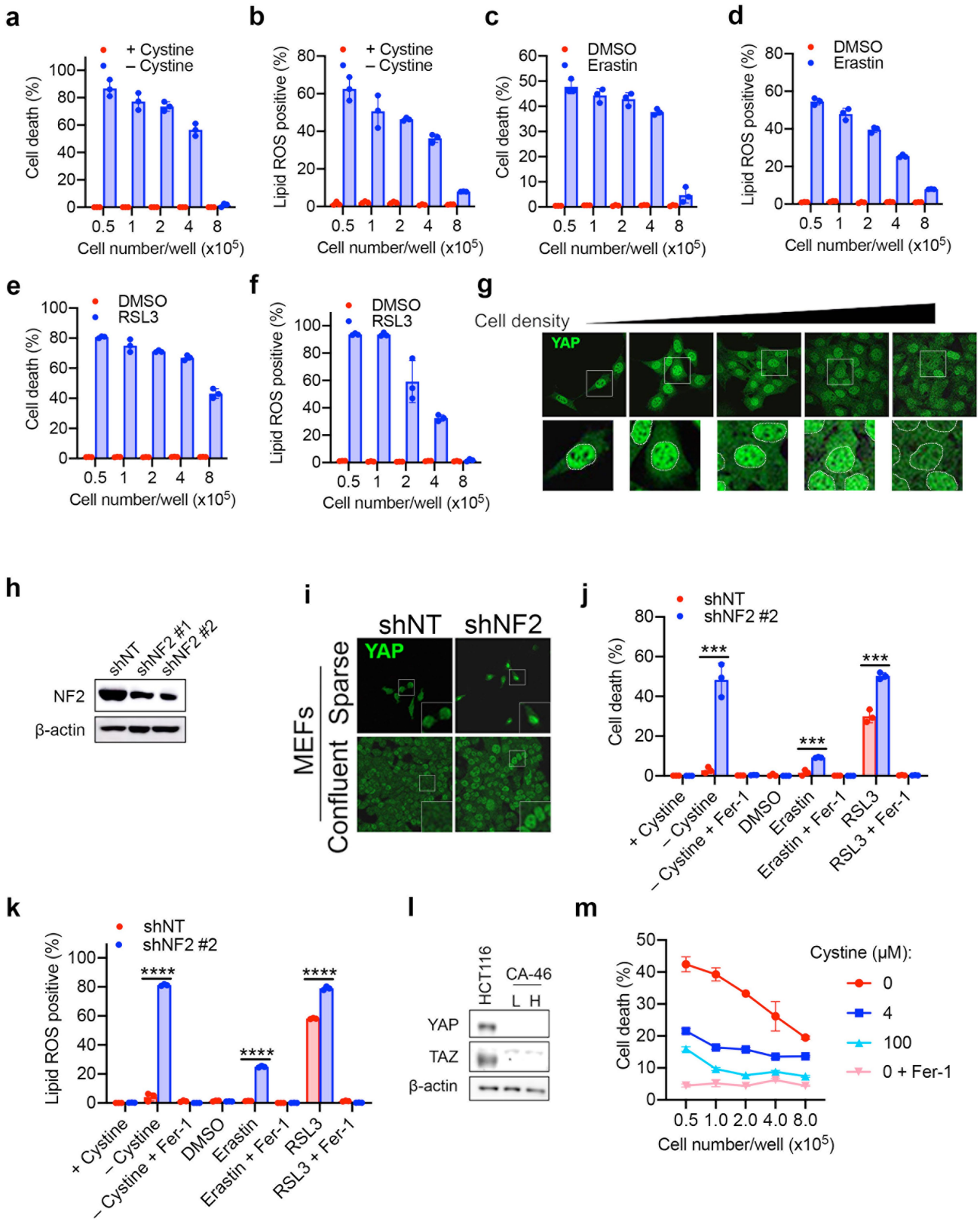
Extended Data Fig. 4 | The Hippo pathway links cell density and intercellular contact to ferroptosis. a, Confluent cells were subjected to cystine starvation for 30 h. Cell death was determined by propidium iodide staining. **b**, HCT116 cells expressing shNT, shECAD, shNF2 or shLATS1/2 as indicated were treated with 5 μM RSL3 with or without 2 μM Fer-1. Cell death was measured at 18 h. $^{****}P < 0.0001$; one-way ANOVA. **c**, Lipid ROS production of cells as in **b** was assessed at 12 h after treatment. $^{****}P < 0.0001$; one-way ANOVA. **d**, Cumulative cell growth curve expressed as the total cell count of HCT116 cells transfected with shNT, shECAD, shNF2 or shLATS1/2. n.s., $P = 0.9497$; two-way ANOVA. **e**, Western blot analysis of the expression and phosphorylation of NF2 in HCT116 cells transfected with enhanced green fluorescent protein (eGFP)-tagged PAK containing a prenylation (CAAX) motif (thus constitutively active), or an inactive mutant form of PAK (K298R). **f**, YAP/TEAD transcriptional activity was measured by a luminescence assay in HCT116 cells expressing activated or inactive PAK and transfected with the 8xGTIIC-luciferase reporter. $^{****}P < 0.0001$; one-way ANOVA. **g**, Cell death was measured in HCT116 cells plated at high density expressing activated or inactive PAK, and treated with cystine-free medium with or without 1 μM Fer-1 for 30 h. $^{****}P < 0.0001$; one-way ANOVA. **h**, Cells were prepared as in **g** and treated with DMSO or 5 μM RSL3 with or without 1 μM Fer-1. Cell death was measured at 24 h. $^{****}P < 0.0001$; one-way ANOVA. All data are mean ± s.d. from $n = 3$ biological replicates.



Extended Data Fig. 5 | NF2 expression correlates with sensitivity to ferroptosis in mesothelioma cell lines. a, Western blot analysis of the expression of LATS1 and LATS2 in the indicated mesothelioma cell lines. **b**, Spheroids were treated with 10 μM erastin for 24 h before SYTOX Green staining. Original magnification, ×40. **c**, Western blot analysis confirming knockdown efficiency of *NF2* shRNA in 211H cells. **d**, Confluent 211H cells transfected with shNT or shNF2 were treated with 1 μM RSL3, with or without 2 μM Fer-1. Cell death (left, 24 h after treatment) and lipid ROS production (right, 16 h) were measured. $^{****}P < 0.0001$; one-way ANOVA. **e**, NF2-mutant Meso33 cells were reconstituted with wild-type NF2, and the expression of NF2 was confirmed by western blot. **f**, Localization of YAP (green) under sparse or confluent conditions in Meso33 cells expressing wild-type NF2 was determined by immunofluorescence. Original magnification, ×200. **g**, Meso33 cells expressing wild-type NF2 were cultured under sparse or confluent conditions and stimulated with cystine-free medium. Cell death was measured by SYTOX Green staining coupled with flow cytometry after 24 h of treatment. Original magnification, ×100. n.s., $P = 0.1874$. $^{*}P = 0.0104$; two-tailed *t*-test. **h**, Meso33 cells expressing wild-type NF2 were cultured as in **g** and the production of lipid ROS was measured after cystine starvation for 16 h. n.s., $P = 0.4860$. $^{*}P = 0.0201$; two-tailed *t*-test. **i**, Meso33 spheroids containing Dox-inducible NF2 were grown in the presence or absence of 1 μg ml^{-1} Dox for 72 h, at which point 10 μM erastin was added. Cell death was measured after 24 h by SYTOX Green staining of spheroids. Original magnification, ×100. All data are mean ± s.d. from $n = 3$ biological replicates.

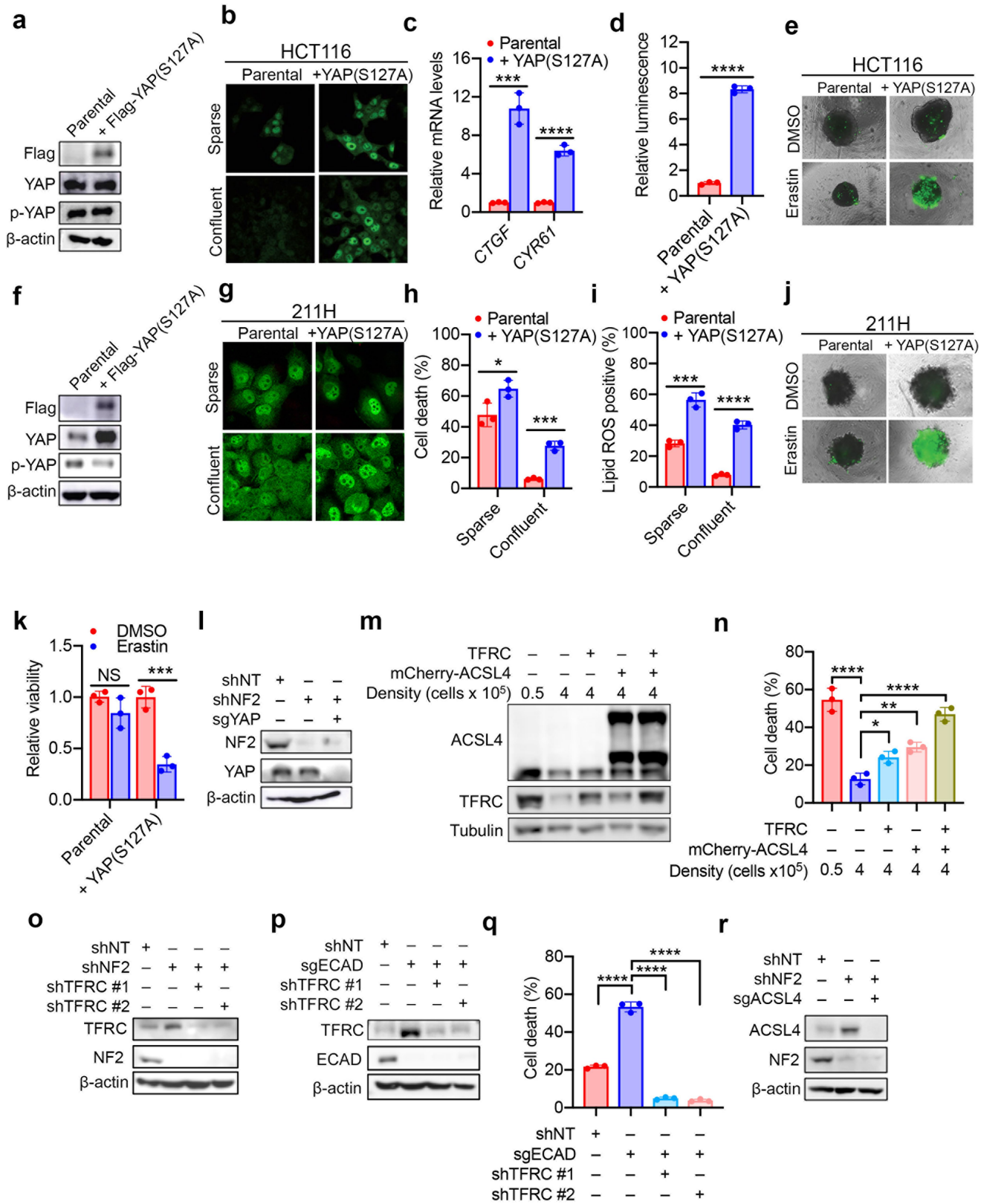


Extended Data Fig. 6 | NCAD suppresses ferroptosis in MSTO-211H cells in a density-dependent manner. a, Western blot analysis of the levels of NCAD, p-YAP and total YAP in 211H cells cultured at different cell densities. **b**, Knockdown efficiency of *NCAD* shRNA (shNcad #1 and #2) was assessed by western blot analysis in 211H cells infected with lentiviruses expressing shNCAD. **c**, Confluent or sparse shNT or shNCAD 211H cells, as indicated, were subjected to cystine starvation for 24 h, at which point cell death was monitored by SYTOX Green staining. Original magnification, ×100. **d**, Flow cytometric quantification of cell death in **c**. n.s., $P = 0.8426$. $^{***}P = 0.0056$, 0.0015 (left to right), $^{****}P < 0.0001$; one-way ANOVA. **e**, Confluent or sparse shNT or shNCAD 211H cells, as indicated, were treated with 1 μM RSL3 for 16 h, at which point cell death was measured by SYTOX Green staining followed by flow cytometry. n.s., $P = 0.3012$, $^{*}P = 0.0315$, $^{****}P < 0.0001$; one-way ANOVA. **f**, Spheroids generated from shNT and shNCAD 211H cells were treated with 10 μM erastin for 24 h, and cell death was determined by SYTOX Green staining. Original magnification, ×40. **g**, Cell viability of spheroids described in **f** was assayed by measuring cellular ATP levels. n.s., $P = 0.4365$, $^{***}P = 0.0006$, $^{****}P < 0.0001$; two-tailed *t*-test. **h**, shNT or shNCAD 211H cells were plated at high density and YAP localization was assessed by immunofluorescence. Original magnification, ×400. **i**, Transcription levels of *CTGF* and *CYR61* measured by qPCR 211H cells plated at high density and transfected with shNT or shNCAD. $^{*}P = 0.0108$, $^{**}P = 0.0016$, $^{***}P = 0.0007$, $^{****}P < 0.0001$; one-way ANOVA. **j**, YAP/TEAD transcriptional activity in 211H cells transfected with shNT or shNCAD was measured by a luciferase assay using the 8xGTIIC-luciferase reporter. $^{***}P = 0.0002$, $^{****}P < 0.0001$; one-way ANOVA. All data are mean ± s.d. from $n = 3$ biological replicates.



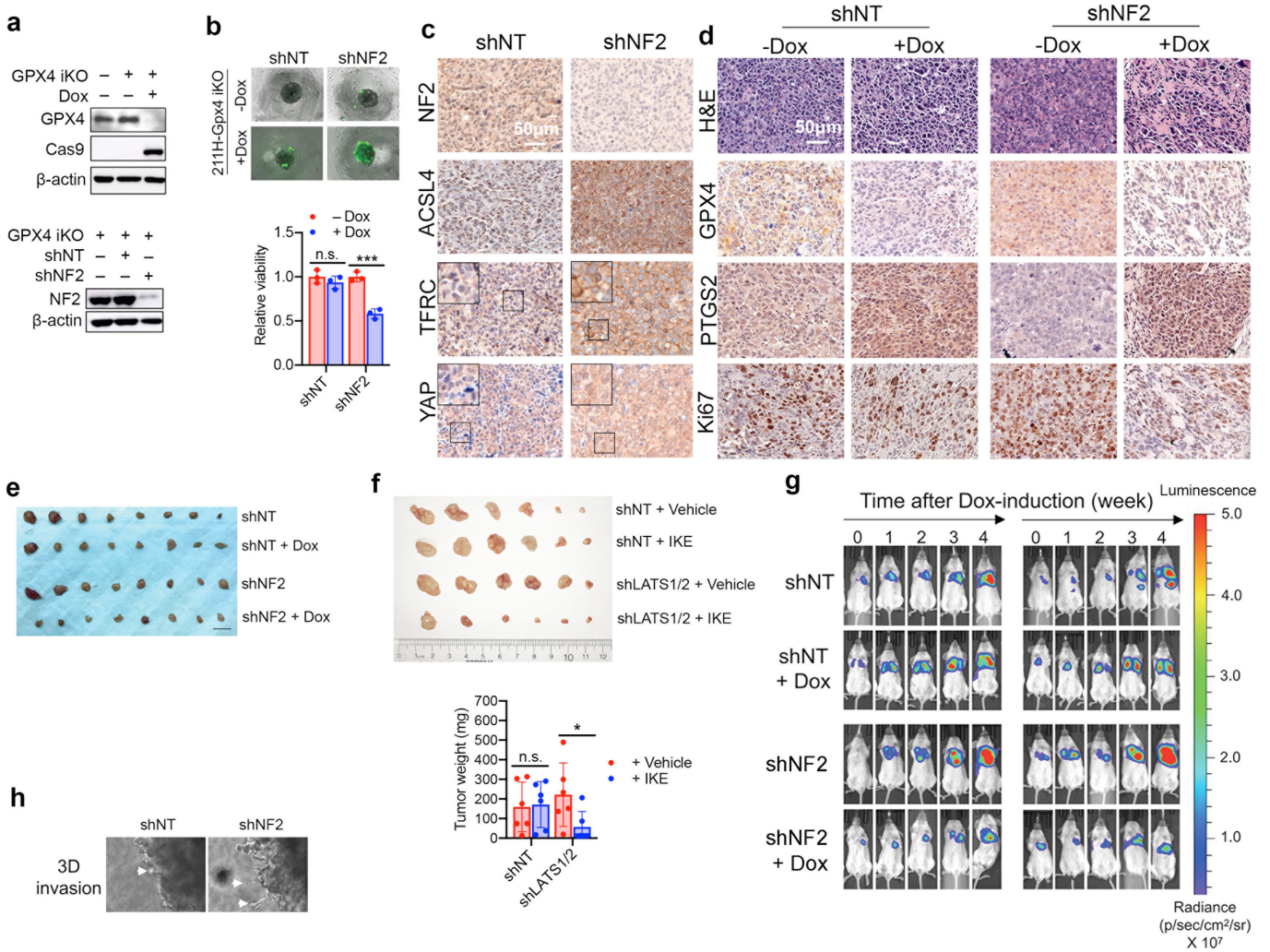
Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Ferroptosis can be regulated by the Hippo pathway in non-epithelial cells. a, Cell death was measured in MEFs after cystine starvation for 12 h. **b**, Cells were treated as in **a** and lipid ROS production was measured at 8 h. **c**, Cell death was measured in MEFs seeded at the indicated densities and treated with 1 μM erastin for 12 h. **d**, Cells were treated as in **c** and production of lipid ROS was measured at 8 h. **e**, MEFs treated with 1 μM RSL3 at the indicated densities were measured for cell death at 8 h. **f**, Cells were treated as in **e** and the production of lipid ROS was measured at 5 h. **g**, Immunofluorescence probing for YAP localization in MEFs seeded at increasing density. Bottom images are enlarged to show localization. Original magnification, ×400. **h**, MEFs were transfected with *NF2* shRNAs (shNF2 #1 and #2), and knockdown efficiency was assessed by western blot. **i**, Immunofluorescence probing for YAP localization after NF2 knockdown in MEFs. Original magnification, ×200. **j**, Increased cell death occurred in confluent MEFs after NF2 depletion (shNF2 #2) and cystine starvation, or treatment with erastin (1 μM, 12 h) or RSL3 (1 μM, 8 h), and this increase was blocked by Fer-1 (2 μM). ***$P = 0.0007$, 0.0007, 0.0006 (left to right); two-tailed *t*-test. **k**, Cells were treated as in **j** and lipid ROS production was assessed at 8 h (cystine starvation, erastin) or 5 h (RSL3). ****$P < 0.0001$; two-tailed *t*-test. **l**, Western blot analysis of expression of YAP and TAZ in CA-46 Burkitt lymphoma cells. **m**, Cell death measurement of CA-46 cells treated as indicated after 24 h. All data are mean ± s.d. from $n = 3$ biological replicates.

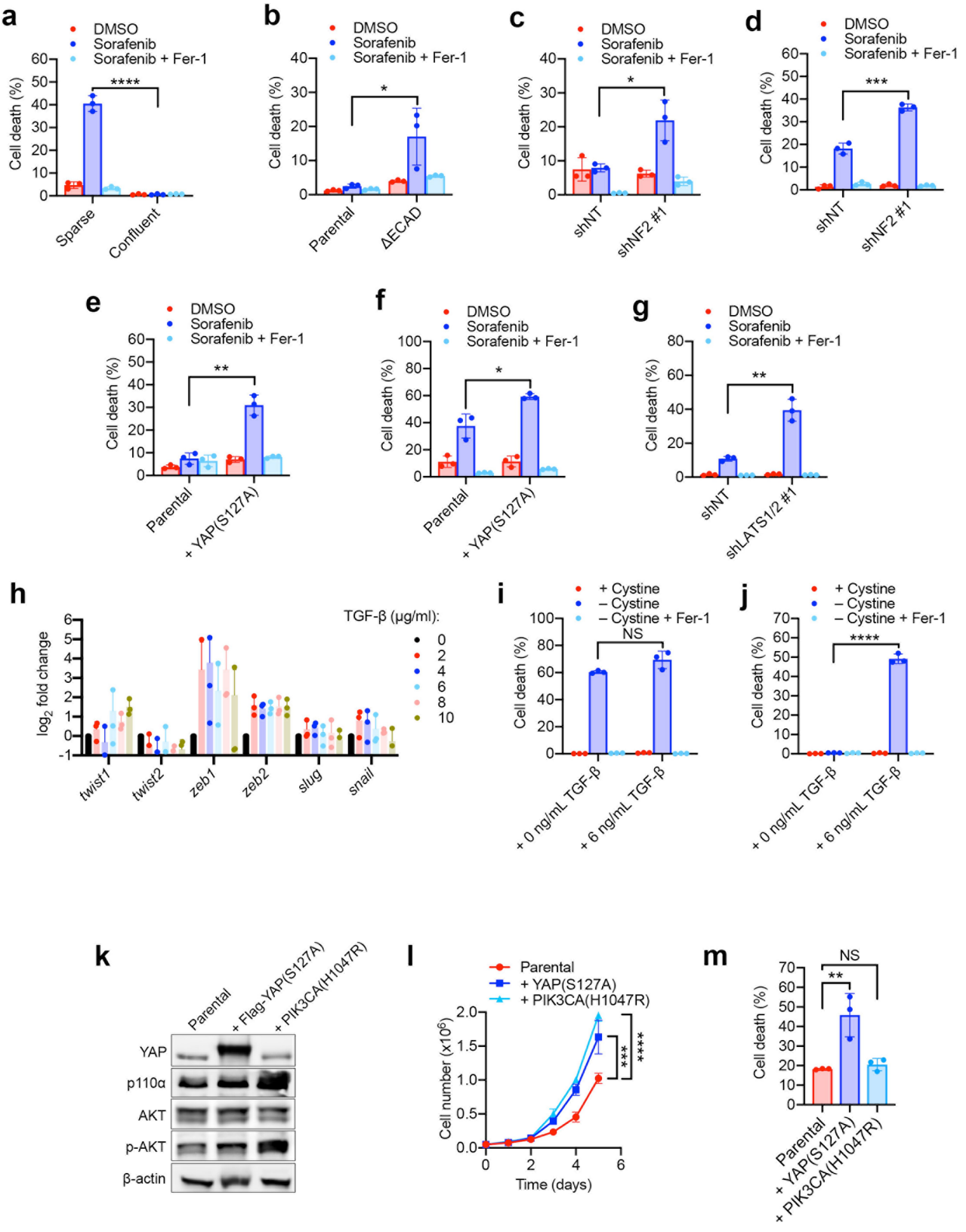


Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | YAP regulates ferroptosis. a, Western blot confirming expression of YAP(S127A) in HCT116 cells. **b**, YAP localization in HCT116 cells expressing YAP(S127A) assessed by immunofluorescence. Original magnification, ×200. **c**, Transcriptional levels of *CTGF* and *CYR61* measured by qPCR in HCT116 cells expressing YAP(S127A). ***$P = 0.0005$, ****$P < 0.0001$; two-tailed *t*-test. **d**, YAP/TEAD transcriptional activity in HCT116 cells expressing YAP(S127A) was measured by a luminescence assay using the 8xGTIIC-luciferase reporter. ****$P < 0.0001$; two-tailed *t*-test. **e**, Spheroids generated from parental and YAP(S127A)-overexpressing HCT116 cells were treated with 15 μM erastin for 30 h, followed by SYTOX Green staining. Original magnification, ×40. **f**, 211H cells were infected with retroviral vectors encoding the Flag–YAP(S127A) mutant, and levels of Flag–YAP, YAP and p-YAP were analysed by western blot. **g**, Localization of YAP (green) was determined by immunofluorescence in 211H cells expressing constitutively active YAP. Original magnification, ×200. **h**, Parental and YAP(S127A)-overexpressing 211H cells were cultured under sparse or confluent conditions and cell death was measured after cystine starvation for 24 h. *$P = 0.0354$, ***$P = 0.0003$; two-tailed *t*-test. **i**, Cells were cultured as in **h** and the production of lipid ROS was measured after 16 h of cystine starvation. ***$P = 0.0006$, *** $P < 0.0001$; two-tailed *t*-test. **j**, Spheroids generated from parental and YAP(S127A)-overexpressing 211H cells were treated with 10 μM erastin for 24 h and cell death was measured by SYTOX staining. Original magnification, ×40. **k**, Cells were cultured as in **j** and cell viability within spheroids was determined by measuring cellular ATP levels. n.s., $P = 0.1534$. ***$P = 0.0009$l; two-tailed *t*-test. **l**, YAP was knocked out by CRISPR–Cas9 (sgYAP) in shNF2 HCT116 cells. **m**, HCT116 cells were transduced with retroviral particles containing mCherry–ACSL4 and/or transfected with TFRC. Expression levels were assayed by western blot. Two bands were detected for mCherry–ACSL4, representing the full-length mCherry–ACSL4 and that with the mCherry tag truncated. **n**, HCT116 cells treated as in **m** were plated at the indicated density and treated with 2 μM RSL3 for 24 h. Cell death was measured by SYTOX Green staining coupled with flow cytometry. *$P = 0.0158$, **$P = 0.0012$, ****$P < 0.0001$; one-way ANOVA. **o**, Western blot analysis confirming knockdown of TFRC in HCT116 shNF2 cells. **p**, Western blot confirming knockdown of TFRC in HCT116 ΔECAD (sgECAD) cells. **q**, Cells as in **p** were treated with medium lacking cystine for 30 h. Cell death was measured by SYTOX Green staining coupled with flow cytometry. ****$P < 0.0001$; one-way ANOVA. **r**, Western blot analysis of HCT116 cells after CRISPR–Cas9-mediated knockout of ACSL4 (sgACSL4) and/or transfection with shNF2. All data are mean ± s.d. from $n = 3$ biological replicates.



Extended Data Fig. 9 | NF2, LATS1 and LATS2 regulate cancer cell sensitivity to ferroptosis in vivo. a, Top, western blot analysis confirming knockout of GPX4 (GPX4 iKO) in 211H cells after treatment with 1 μg ml^{-1} Dox for 5 days (top). Bottom, cells were infected with the indicated hairpins. **b**, Spheroids formed by shNT-GPX4-iKO or shNF2-GPX4-iKO 211H cells were treated with or without Dox for five days. Cell death and viability, respectively, were determined by SYTOX staining (top) and cellular ATP levels (bottom). Original magnification, ×40. n.s., $P = 0.3523$, $^{***}P = 0.0007$; two-tailed t-test. **c**, shNT-GPX4-iKO and shNF2-GPX4-iKO cells were subcutaneously injected into nude mice. The effect of NF2 knockdown on xenografted tumours was validated by immunostaining of NF2, ACSL4, TFRC and YAP, all counterstained with haematoxylin (blue), on sections of tumours bearing shNT and shNF2 as indicated. Inset images are enlarged to show TFRC expression at plasma membranes and increase in nuclear localization of YAP. **d**, shNT-GPX4-iKO cells and shNF2-GPX4-iKO cells were subcutaneously injected into nude mice followed by treatment with or without Dox (to induce GPX4 knockout; $n = 8$ mice per group). Representative haematoxylin and eosin (H&E) and immunostaining images of GPX4, PTGS2 and Ki67, all counterstained with haematoxylin (blue), are shown from sections of xenografted tumours. **e**, Images of resected MSTO-211H subcutaneous tumours. Scale bar, 1 cm. **f**, Mice injected subcutaneously with HCT116 cells expressing shNT or shLATS1/2 as indicated were treated with or without IKE. Top, images of resected HCT116 shNT or shLATS1/2 tumours. Bottom, mass of resected tumours. $n = 6$ mice per group. n.s., $P = 0.8677$. $^{*}P = 0.0486$; two-tailed t-test. **g**, Representative BLI showing the tumour growth of indicated cells in an orthotopic model of mesothelioma in nonobese diabetic/severe combined immunodeficiency (NOD/SCID) mice. Dox treatment started when the average total flux reached 10^8 photons per second (time point 0). **h**, Tumour spheroids of 211H cells expressing shNT or shNF2 were grown in Matrigel, and invasion was monitored. In the representative images, arrows show protrusions extruded from the main body of spheroids. Original magnification, ×400. All data are mean ± s.d. from $n = 3$ biological replicates.



Extended Data Fig. 10 | See next page for caption.

Extended Data Fig. 10 | The Hippo pathway as a potential biomarker for predicting cancer cell sensitivity to ferroptosis. a, Cell death was measured in HCT116 cells seeded at 0.5×10^5 cells per 3.5 cm^2 well (sparse) or 4×10^5 cells per 3.5 cm^2 well (confluent) and grown for 24 h. Cells were treated with DMSO, 10 μM sorafenib or 10 μM sorafenib plus 2 μM Fer-1 as indicated for 24 h. ****$P < 0.0001$; two-tailed *t*-test. **b**, Cell death was measured in parental or ΔECAD HCT116 cells seeded at 4×10^5 cells per 3.5 cm^2 well and grown for 24 h. Cells were treated as in **a**. *$P = 0.0394$; two-tailed *t*-test. **c**, **d**, Cell death was measured in HCT116 (**c**) or 211H (**d**) shNT or shNF2 cells seeded at high density and treated as in **a**. *$P = 0.0167$, ***$P = 0.0004$; two-tailed *t*-test. **e**, **f**, Cell death was measured in HCT116 (**e**) or 2211H (**f**) cells expressing parental or YAP(S127A) cells seeded at high density and treated as in **a**. *$P = 0.0143$, **$P = 0.0014$; two-tailed *t*-test. **g**, Cell death was measured in HCT116 shNT or shLATS1/2 cells seeded at high density and treated as in **a**. **$P = 0.0017$; two-tailed *t*-test. **h**, NF639 cells, derived from mouse mammary tumours containing MMTV-neu, were treated with various concentrations of TGFβ for 48 h. mRNA expression of a panel of EMT-related genes was assayed by qPCR. **i**, NF639 cells were treated with or without 6 ng μl^{-1} TGFβ for 48 h, at which point they were plated at low density (0.8×10^5 cells per 3.5 cm^2 well), grown overnight and treated with medium containing or lacking cystine, with or without 1 μM Fer-1 for 12 h, followed by cell death measurement. n.s., $P = 0.0777$; two-tailed *t*-test. **j**, NF639 cells were plated at 3.2×10^5 cells per 3.5 cm^2 well, grown overnight and treated as described in **a**. ****$P < 0.0001$; two-tailed *t*-test. **k**, 211H cells were infected with YAP(S127A) or the activated mutant PIK3CA(H1047R). Lysates were probed for overexpression and phosphorylated AKT (p-AKT; S473) to confirm the activity of PIK3CA(H1047R). **l**, Approximately 50,000 211H cells were seeded in 3.5-cm^2 plates and grown for 5 days. Cells were counted daily. ***$P = 0.0007$, ****$P < 0.0001$; two-way ANOVA. **m**, Cell death was measured by flow cytometry in 211H cells seeded at high density (8×10^5 cells per 3.5-cm^2 well) after cystine starvation for 24 h. n.s., $P = 0.8838$. **$P = 0.0041$; one-way ANOVA. All data are mean ± s.d.; $n = 3$ biological replicates.

nature research

Corresponding author(s): Xuejun Jiang, Zhi-Nan Chen

Last updated by author(s): May 8, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	BD FACSDiva (v8.0.1) was used to collect flow cytometry data. Tecan iControl software (v1.12) was used for the collection of viability/luminescence data in cultured cells. In vivo imaging studies were performed using Caliper Life Sciences Living Image software (v4.3.1).
Data analysis	BD FACSDiva (v8.0.1) and FlowJo (v10.0.7) were used for analysis of flow cytometry data. GraphPad Prism (v7.0) was used for statistical analysis of all data.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All data generated or analysed during this study are included in the manuscript and supplementary information files.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical calculations were performed to predetermine sample size.
Data exclusions	No data has been excluded from this study.
Replication	All experiments presented in this study yielded reproducible results for a minimum of three independent replicates.
Randomization	For animal experiments, animals, were randomized for testing different conditions.
Blinding	Animal studies were blinded for the collection of measurements.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Western blots: GPX4 (Abcam, ab125066, GR251529-31, [EPNCIR144], 1:1000); E-cadherin (Abcam, ab76055, GR317373-2, [M168], 1:1000); N-cadherin (Abcam, ab76057, GR259225-14, polyclonal, 1:1000); pan-cadherin (Cell Signaling, 4086T, 2, polyclonal, 1:1000); Merlin (Cell Signaling, 12888S, 1, [D3S3W], 1:1000); pMerlin-Ser518 (Cell Signaling, 13281, 1, [D5A4I], 1:1000); LATS1 (Cell Signaling, 3477T, 7, [C66B5], 1:1000); LATS2 (Cell Signaling, 5888S, [D83D6], 1:1000); YAP (Cell Signaling, 14074S, 2, [D8H1X], 1:1000); pYAP-Ser127 (Cell Signaling, 4911S, 5, polyclonal, 1:1000); ACSL4 (Thermo Fisher, PA5-27137, polyclonal, 1:1000); transferrin receptor (Abcam, ab214039, GR189553-8, polyclonal, 1:1000); TAZ (Cell Signaling, 83669S, 1, [E8E9G], 1:1000); p110 alpha (Cell Signaling, 4249S, 7, [C73F8], 1:1000); AKT (Cell Signaling (2920S, 16, [40D4], 1:1000); pAKT-Ser473 (Cell Signaling, 4060S, 3, D9E, 1:1000); Cas9 (Cell Signaling, 14697S, 3, [7A9-3A3], 1:1000); HA (Sigma-Aldrich, H3663, [HA-7], 1:2000); Flag (Sigma-Aldrich, F1804, SLBW3851, [M2], 1:2000); GFP (Invitrogen, A11122, 1024102, polyclonal, 1:3000); alpha tubulin (Calbiochem, CP06, D00175772, [DM1A], 1:3000); beta actin (Sigma-Aldrich, A1978, [AC-15], 1:3000); rabbit IgG-HRP (Thermo Fisher, 31458, polyclonal, 1:10000); mouse IgG-HRP (Thermo Fisher, 31430, polyclonal, 1:10000). Immunofluorescence: E-cadherin (Abcam, ab76055, GR317373-2, [M168], 1:200); YAP (Cell Signaling, 14074S, 2, [D8H1X], 1:200); rabbit IgG-AlexaFluor488 (Invitrogen, A11008, 34732A, polyclonal, 1:500); rabbit IgG-AlexaFluor594 (Invitrogen, A11012, 1810936, polyclonal, 1:500); mouse IgG-AlexaFluor488 (Invitrogen, A11029, 673781, polyclonal, 1:500); mouse IgG-AlexaFluor594 (Invitrogen, A11005, 610868, polyclonal, 1:500). Immunohistochemistry: GPX4 (Abcam, ab125066, GR251529-31, [EPNCIR144], 1:100); YAP (Cell Signaling, 14074S, 2, [D8H1X], 1:400); ACSL4 (Thermo Fisher, PA5-27137, polyclonal, 1:200); transferrin receptor (Abcam, ab214039, GR189553-8, polyclonal, 1:400); Merlin (Abcam, ab88957, GR310755-14, [AF1G4], 1:100); PTGS2 (Cell Signaling, 12282, 4, [D5H5], 1:400); Ki67 (Cell Signaling, 9449, 4, [8D5], 1:400). ChIP: TEAD4 (Abcam, ab58310, GR3205108-1, 10 µg/sample)
Validation	Antibodies were only used for applications and organisms verified by the manufacturer.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	MEF, HEK293T, HepG2, H1650, BT474, MDA-MB-231, NF639, and CA-46 cells were acquired from ATCC. PC9 cells were acquired from Sigma-Aldrich. The human mesothelioma cell lines 211H, H2452, H28, H-meso, Meso33, Meso9, Meso37,

	H2052, JMN, and VAMT were a gift from the lab of Filippo Giancotti (MD Anderson, Houston, TX). 211H, H2452, H28, and H2052 are available on ATCC.
Authentication	All ATCC cell lines were authenticated by STR DNA profiling analysis through ATCC. All other cell lines were also profiled through STR analysis and found to not match any cell line available from ATCC.
Mycoplasma contamination	All cell lines tested negative for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No cell lines used in this study are listed in the ICLAC database.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Athymic nude mice: female, 6-8 weeks. Nod/SCID mice: female, 6-8 weeks.
Wild animals	This study did not involve wild animals.
Field-collected samples	This study did not involve field-collected samples.
Ethics oversight	All protocols used in this study were approved by the Memorial Sloan Kettering Institutional Animal Care and Use Committee (IACUC).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	All cells used in Flow Cytometry are cultured cell lines. Cell death was analyzed by propidium iodide (Invitrogen, Waltham, MA, USA) or SYTOX Green (Invitrogen) staining followed flow cytometry. Samples were washed with PBS before flow cytometry. To analyze lipid peroxidation, cells were stained 5 µM BODIPY-C11 (Invitrogen) for 30 minutes at 37°C followed by flow cytometric analysis.
Instrument	LSRII (BD Biosciences)
Software	BD FACSDiva (v8.0.1) was used to collect data, BD FACSDiva (v8.0.1) and Flowjo (v10.0.7) were used to analyze the data.
Cell population abundance	For all samples, a minimum of 10,000 cells were collected.
Gating strategy	For cell death detection, unstained controls (PBS alone without staining) were used as negative control. The first plot (FSC-A vs. SSC-A) was drawn to exclude debris as they tend to have lower forward scatter levels. The second and third plots (SSC-W vs. SSC-H and FSC-W vs. FSC-H) are drawn to remove doublets from the analysis. Then a single parameter histogram was used to identify cells which would be positive for the dye. A gate boundary is made based on the unstained cells and the peaks of the histogram. For lipid ROS staining, samples without BODIPY-C11 was used as negative control. A figure exemplifying the gating strategy are provided in Supplementary information.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

CORRECTION

https://doi.org/10.1038/s41586-019-1455-1

Author Correction: Epigenetic stress responses induce muscle stem-cell ageing by *Hoxa9* developmental signals

Simon Schwörer, Friedrich Becker, Christian Feller, Ali H. Baig, Ute Köber, Henriette Henze, Johann M. Kraus, Beibei Xin, André Lechel, Daniel B. Lipka, Christy S. Varghese, Manuel Schmidt, Remo Rohs, Ruedi Aebersold, Kay L. Medina, Hans A. Kestler, Francesco Neri, Julia von Maltzahn, Stefan Tümpel & K. Lenhard Rudolph

Correction to: *Nature* https://doi.org/10.1038/nature20603, published online 30 November 2016.

In this Letter, errors occurred in the following figures. In Extended Data Fig. 6e, the 'shScr, Aged donor' image is a duplicate of the 'Vehicle, Aged donor' image in Fig. 3f. The images in Extended Data Fig. 6e represent differences in engraftment levels under four experimental conditions; however, these reflect the lower end of the observed overall engraftment rate in the experiment. Figure 1 of this Amendment shows the corrected panels for Extended Data Fig. 6e, with images from the original experiment that best reflect the differences in, and the overall level of, the engraftment rates between the conditions under study (the original images from Extended Data Fig. 6e are shown for comparison).

In addition, there are errors in the Source Data for Figs. 3d, 4k, Extended Data Figs. 4f–h, 7f, s, t and 9m–o, q–s due to copy-and-paste errors or due to the presentation of controls that were used for the calculation of *P* values or error bars shown in the figures. One value that was identified as an outlier in Extended Data Fig. 10g was not labelled as such in the original Source Data and was erroneously included for graphical depiction. See Supplementary Information to this Amendment for the corrected Source Data files, and Figs. 2, 3 and 4 of this Amendment for the corrected and original panels for Figs. 3d, 4k and Extended Data Fig. 4f, g, respectively. For the calculation of the *P* value in Extended Data Fig. 6b, we applied a one-sided paired ratio Student's *t*-test (not, as stated, a two-sided Student's *t*-test).

In addition, in Extended Data Fig. 7m, n, p, r, 9i, q–s and 10d, e of the original Letter, errors occurred in data scaling that affect the calculation of *P* values and the graphical presentation of the data. See Figs. 5, 6 and 7 of this Amendment for the corrected and incorrect panels for Extended Data Fig. 7f, m, n, p–t, 9i, m, q–s and 10d, e, g, respectively, and Supplementary Information to this Amendment for the corrected Source Data. The errors in data scaling occurred because two methods of data scaling were used throughout the study. In some experiments, data of the experimental groups were scaled to the average of the control values; in other experiments, data of the experimental groups were scaled to each of the corresponding controls of a biological repeat, set to 1 or 100. Although both methods of scaling are valid, they should not be combined within one experiment, which happened in the aforementioned figures. This has now been corrected and we include a detailed description of our scaling approach in the Supplementary Information to this Amendment.

The outlined corrections do not change the conclusions of the original Letter, and we apologize for any confusion that these errors may have caused. The original Letter has not been corrected.

Supplementary Information is available in the online version of this Amendment.

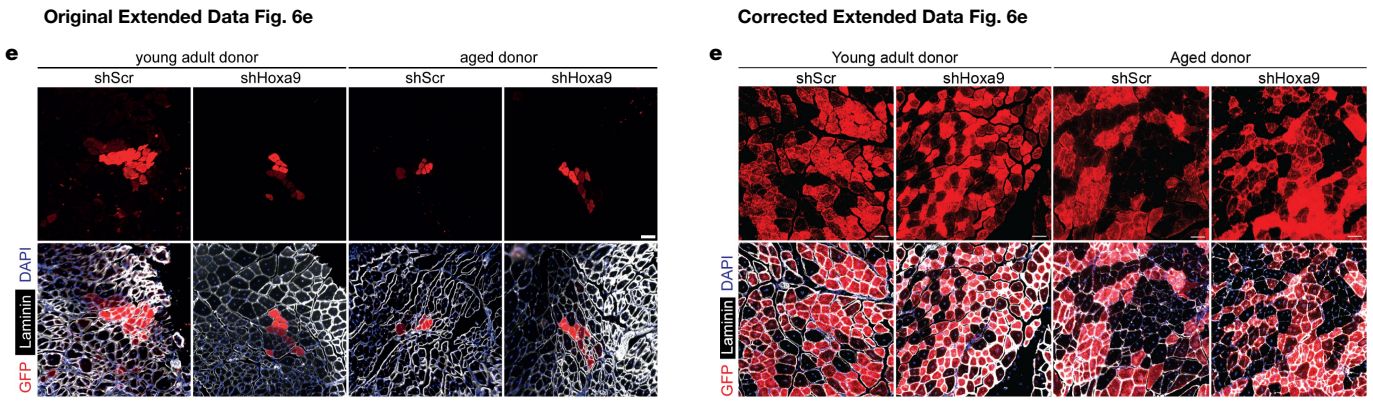


Fig. 1 | This is the corrected Extended Data Fig. 6e (right) and the original Extended Data Fig. 6e (left) published in the original Letter. All images have been replaced in the corrected figure.

CORRECTIONS & AMENDMENTS

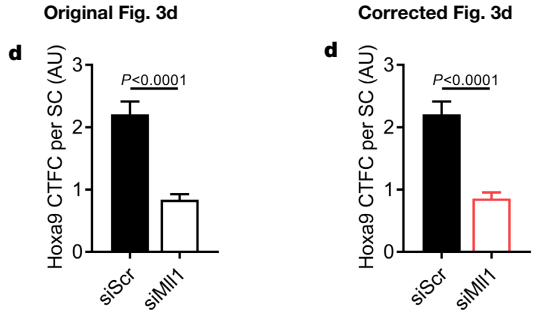


Fig. 2 | This is the corrected Fig. 3d (right) and the original Fig. 3d (left) published in the original Letter. Changes are highlighted in red.

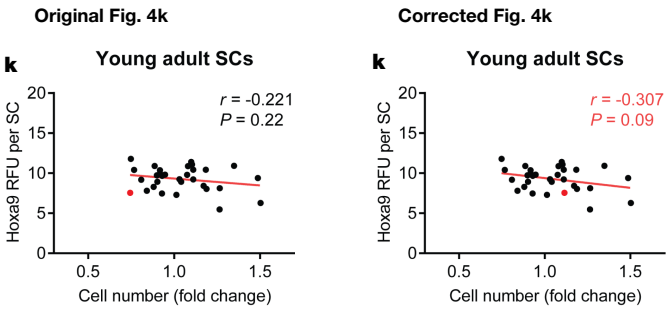


Fig. 3 | This is the corrected Fig. 4k (right) and the original Fig. 4k (left) published in the original Letter. Changes are highlighted in red. The data point that changes is also marked in red in the original figure as well as in the corrected figure.

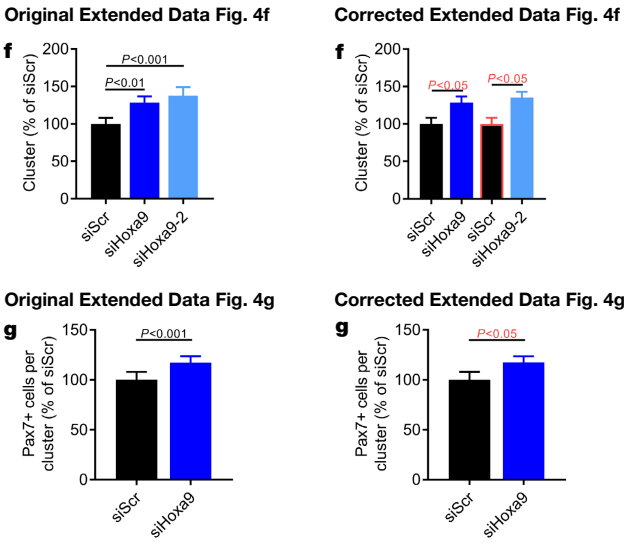


Fig. 4 | This is the corrected Extended Data Fig. 4f, g (right) and the original Extended Data Fig. 4f, g (left) published in the original Letter. Changes are highlighted in red.

CORRECTIONS & AMENDMENTS

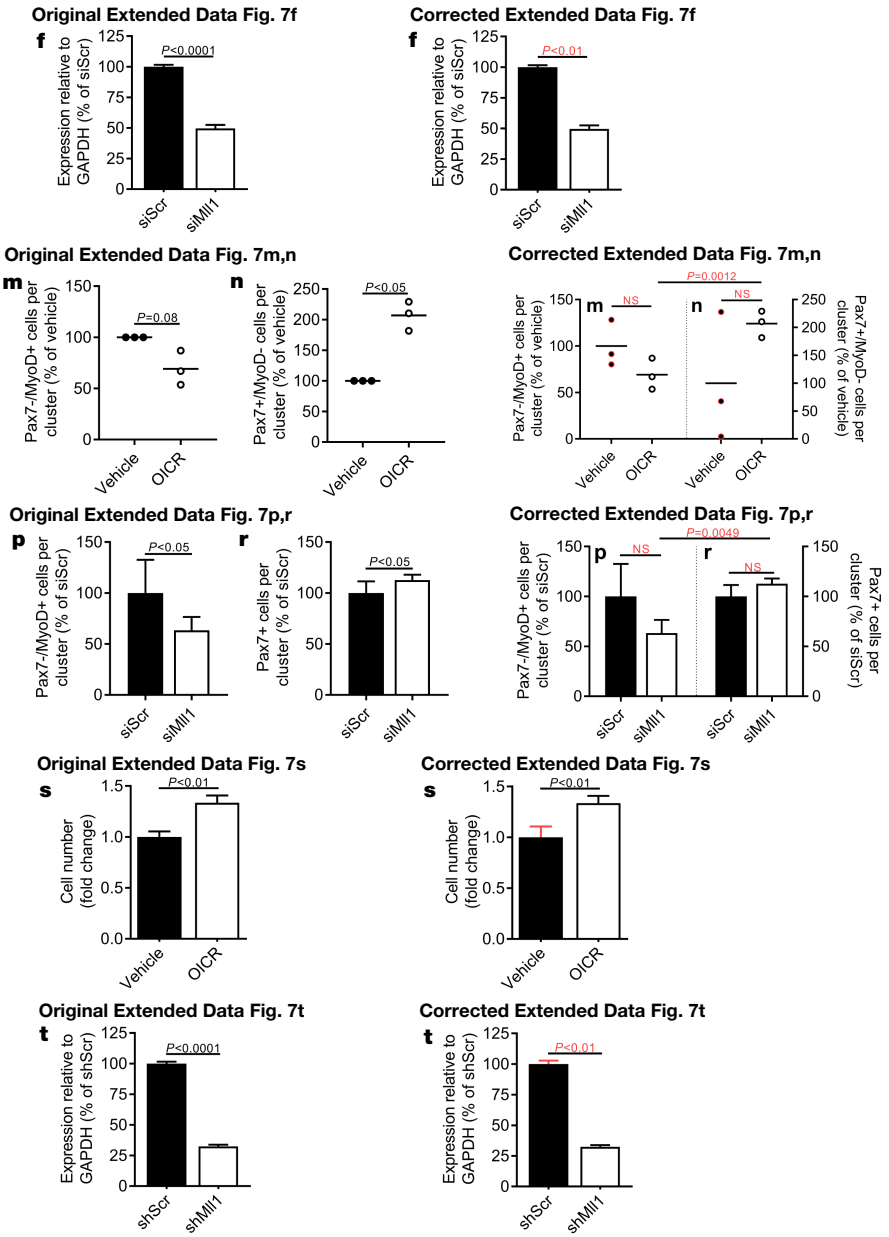


Fig. 5 | This is the corrected Extended Data Fig. 7f, m, n, p, r–t (right) and the original Extended Data Fig. 7f, m, n, p, r–t (left) published in. the original Letter. Changes are highlighted in red.

CORRECTIONS & AMENDMENTS

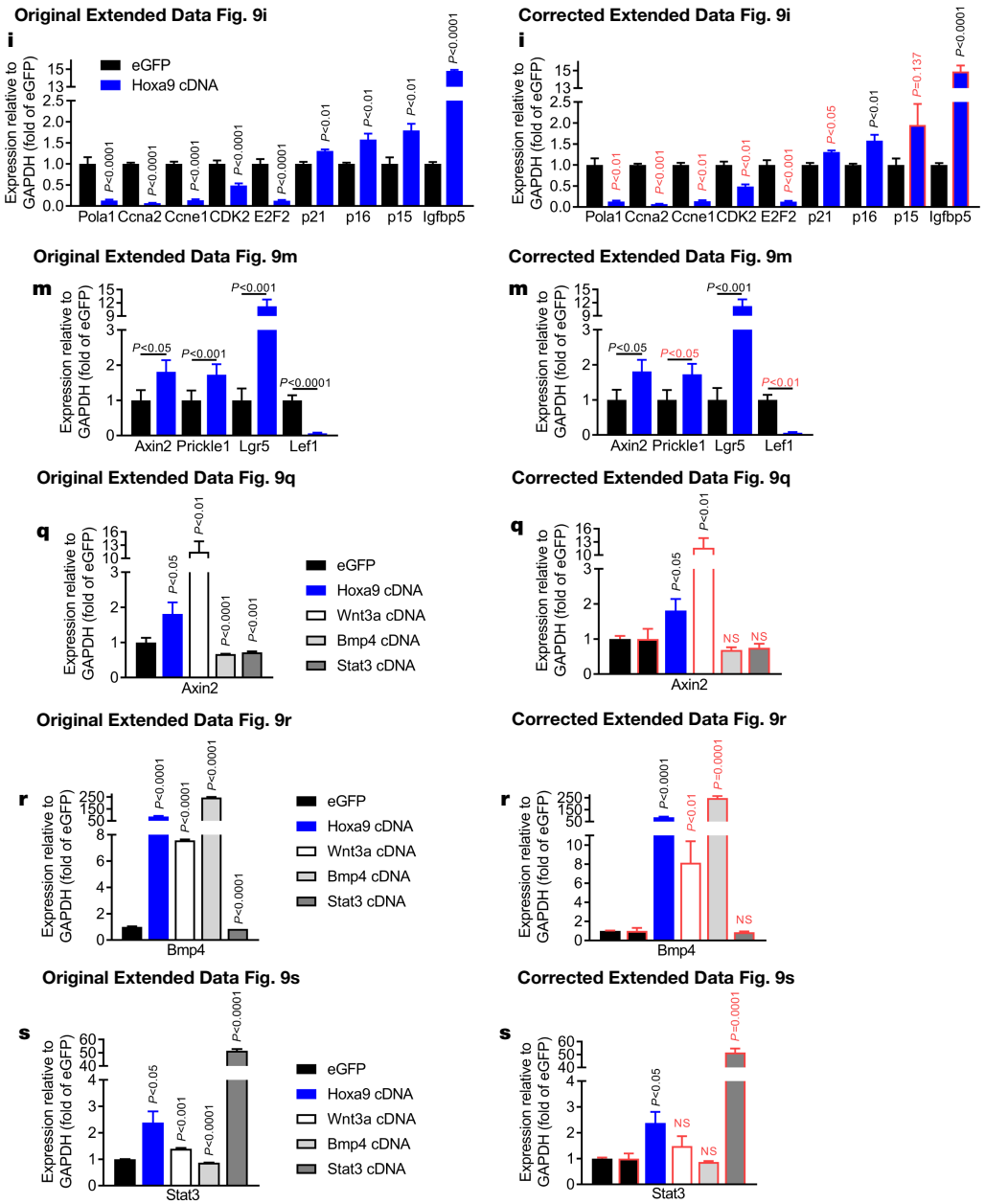


Fig. 6 | This is the corrected Extended Data Fig. 9i, m, q, r, s (right) and the original Extended Data Fig. 9i, m, q, r, s (left) published in the original Letter. Changes are highlighted in red.

CORRECTIONS & AMENDMENTS

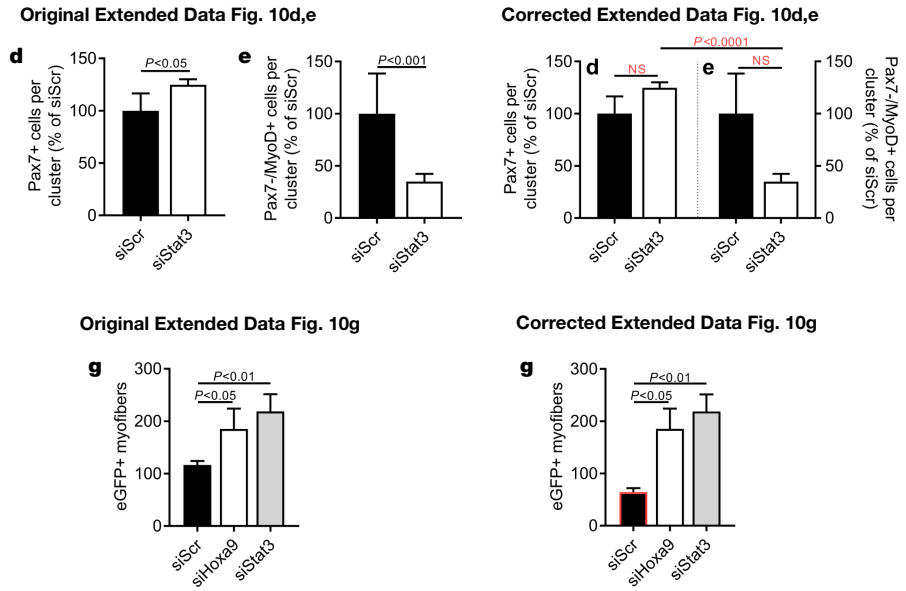


Fig. 7 | Fig. 7 This is the corrected Extended Data Fig. 10d, e, g (right) and the original Extended Data Fig. 10d, e, g (left) published in the original Letter. Changes are highlighted in red.

CORRECTION

https://doi.org/10.1038/s41586-019-1415-9

Author Correction: Advanced maturation of human cardiac tissue grown from pluripotent stem cells

Kacey Ronaldson-Bouchard, Stephen P. Ma, Keith Yeager, Timothy Chen, LouJin Song, Dario Sirabella, Kumi Morikawa, Diogo Teles, Masayuki Yazawa & Gordana Vunjak-Novakovic

Correction to: *Nature* https://doi.org/10.1038/s41586-018-0016-3, published online 04 April 2018.

In this Letter, there are several errors to the following figures. In Fig. 2e, the *y*-axis labels should have been '0, 50, 100, 150, 200, 250 and 300' rather than '0, 50, 100, 100, 200, 200 and 300'. In Extended Data Fig. 5g, owing to an error during figure preparation, the adult human heart tissue group was duplicated and different regions of the same section were shown for both the 'adult' and the 'intensity' images. This figure has been corrected online, and the original, incorrect panels are shown as Fig. 1 of this Amendment for transparency. Raw data for the images in Extended Data Fig. 5g and additional contemporary transmission electron microscopy (TEM) images are included in 'SI ED5.zip' in Figshare (https://doi.org/10.6084/m9.figshare.5765559). In addition, the same TEM micrograph of mitochondria was used in Fig. 2d and the last panel in Extended Data Fig. 6 (labelled 'Mitochondria'), which focused on two different areas to illustrate lipid droplets and the ultrastructure of mitochondria. The same micrograph was shown at different magnifications (scale bars in Fig. 2d and Extended Data Fig. 6 are 200 nm and 500 nm, respectively) and orientations. Raw data for these images and additional contemporary TEM images are included in 'SI ED2' in Figshare (https://doi.org/10.6084/m9.figshare.5765559).

In Extended Data Fig. 8f, g, the boxes indicating the position of the insets were misplaced. This figure has been corrected online, and the original incorrect panels are shown as Fig. 2 to this Amendment for transparency. High-resolution images of Extended Data Fig. 8 and the individual panels are included in 'SI ED8.zip' in Figshare (https://doi.org/10.6084/m9.figshare.5765559). All data were analysed (no exclusions), as specified in the Reporting Summary. In addition, in Fig. 1d and Extended Data Fig. 3f data are 'mean ± s.e.m.' rather than 'mean ± s.d', and the legend and figure, respectively, have been updated accordingly.

In the Methods, the pulsatile electrical field voltage should be '3.5–4.0 V cm^{-1}' rather than '4.5 mV' (as correctly shown in Extended Data Fig. 1e). We also wish to clarify that the immunostainings in Extended Data Fig. 5e, f were performed in two adjacent tissue sections, as specified in the figure legend. Finally, we wish to include two additional references to further document the T-tubule system in adult human heart tissue. These references should appear after the following text: "…dependent on both intracellular calcium reserves and the proximity of $Ca_V1.2$ channels and T-tubules…", and have been included as references 28 and 29, with all subsequent citations renumbered. This Letter has been corrected online.

28. Kaprielian, R. R., Stevenson, S., Rothery, S. M., Cullen, M. J. & Severs, N. J. Distinct patterns of dystrophin organization in myocyte sarcolemma and transverse tubules of normal and diseased human myocardium. *Circulation* **101**, 2586–2594 (2000).
29. Sulkin, M. S. et al. Nanoscale three-dimensional imaging of the human myocyte. *J. Struct. Biol.* **188**, 55–60 (2014).

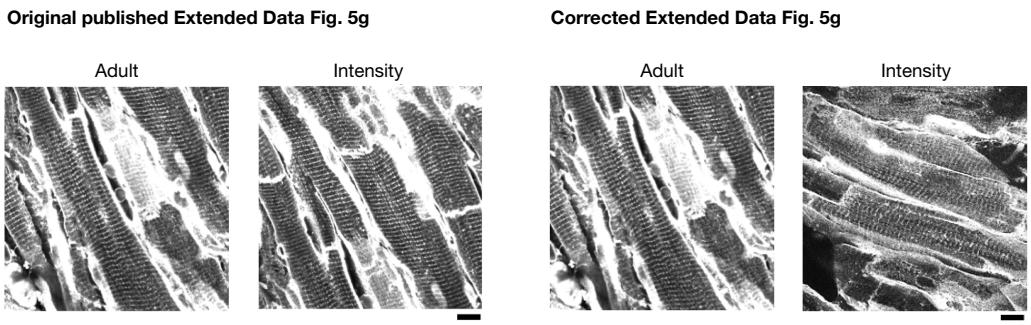


Fig. 1 | This figure shows the original, published 'adult' and 'intensity' images from Extended Data Fig. 5g, and the corrected versions.

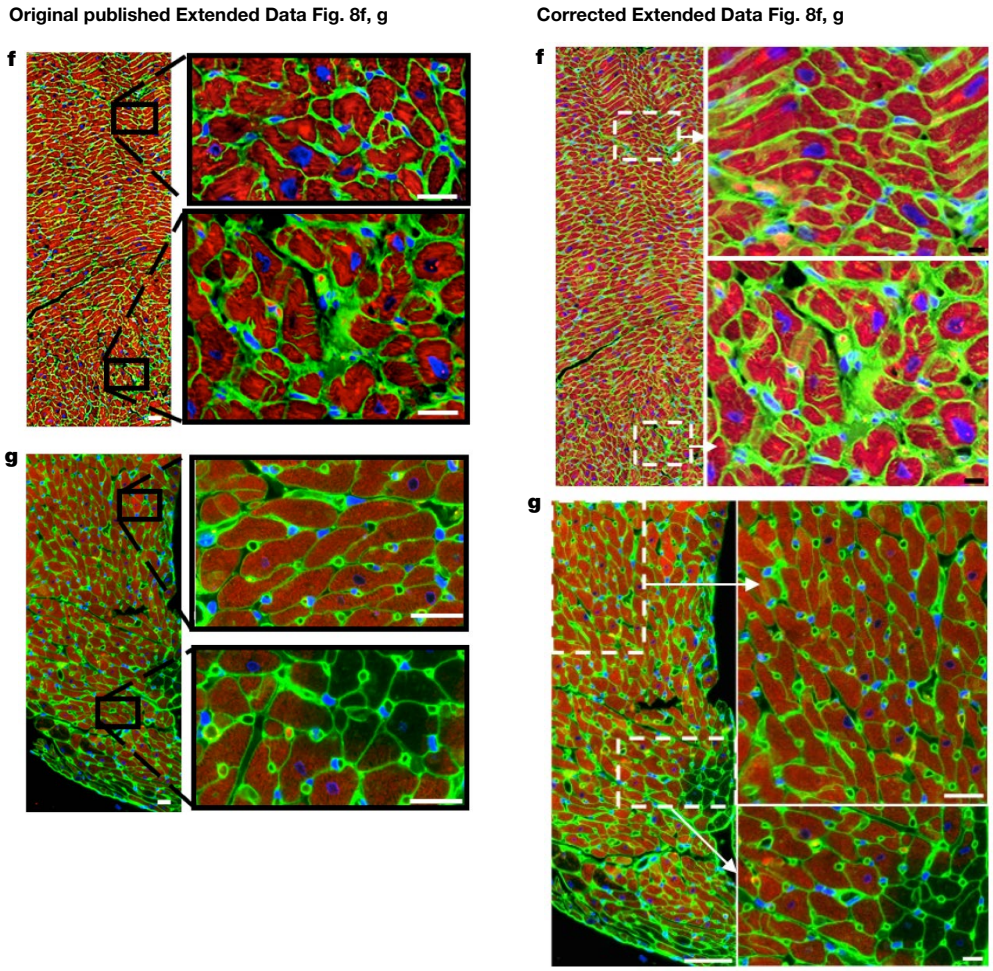


Fig. 2 | This figure shows the original, published Extended Data Fig. 8f, g, and the corrected Extended Data Fig. 8f, g.

CORRECTION

https://doi.org/10.1038/s41586-019-1459-x

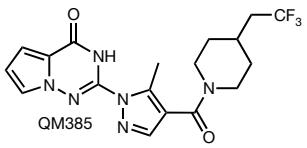
Publisher Correction: The metabolite BH4 controls T cell proliferation in autoimmunity and cancer

Shane J. F. Cronin, Corey Seehus, Adelheid Weidinger, Sebastien Talbot, Sonja Reissig, Markus Seifert, Yann Pierson, Eileen McNeill, Maria Serena Longhi, Bruna Lenfers Turnes, Taras Kreslavsky, Melanie Kogler, David Hoffmann, Melita Ticevic, Débora da Luz Scheffer, Luigi Tortola, Domagoj Cikes, Alexander Jais, Manu Rangachari, Shuan Rao, Magdalena Paolino, Maria Novatchkova, Martin Aichinger, Lee Barrett, Alban Latremoliere, Gerald Wirnsberger, Guenther Lametschwandtner, Meinrad Busslinger, Stephen Zicha, Alexandra Latini, Simon C. Robson, Ari Waisman, Nick Andrews, Michael Costigan, Keith M. Channon, Guenter Weiss, Andrey V. Kozlov, Mark Tebbe, Kai Johnsson, Clifford J. Woolf & Josef M. Penninger

Correction to: *Nature* https://doi.org/10.1038/s41586-018-0701-2, published online 07 November 2018.

In this Letter, owing to an error in the production process, author Martin Aichinger was inadvertently associated with affiliation 14 (Karolinska Institute, Department of Medicine Solna, Center for Molecular Medicine, Karolinska University Hospital Solna, Stockholm, Sweden) instead of affiliation 13 (Research Institute of Molecular Pathology, Vienna Biocenter, Campus-Vienna-Biocenter 1, Vienna, Austria). In addition, the chemical structure of QM385 in Fig. 3a was incorrect. Figure 1 of this Amendment shows the incorrect and corrected structures, for transparency. These errors have been corrected online.

Original Fig. 3a



Corrected Fig. 3a

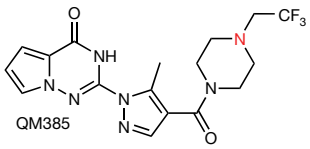


Fig. 1 | This figure shows the incorrect and corrected structures of QM385 in Fig. 3a of the original Letter.

MOST READ Twenty things I wish I'd known when I started my PhD **go.nature.com/2qqw440**

COLLECTION Failure and how it makes better science **go.nature.com/2ms3hhw**

JOBS Check out the latest science-careers listings **naturecareers.com**

ALICE MOLLON/GETTY

COLUMN

Strength in numbers

Share veteran PhD students' experiences with new starters, says **Sarah Masefield**.

When I started my PhD in health sciences in 2016, I knew it was a risk. I had a history of depression, and I thought the programme might trigger a recurrence. What I hadn't expected was the extreme anxiety that I experienced. Over the Christmas holidays of my second year, I woke up every day with my heart racing and feeling sick, knowing that to reach my next deadline I had to spend another day trying to make progress with my systematic-review chapter. My only full day off during that period was Christmas Day.

Instead of seeking help, I stopped communicating with my supervisors because I felt incompetent. I worried that talking to them would expose and shame me more. I was not willing to carry on at the further expense of my health, and of my relationship with my partner. I decided that if something didn't change soon, I'd have to drop out.

Fortunately, I'd made friends with other PhD students in my department at the University of York, UK. We discussed our research projects and shared guidance from our supervisors and other students. Hearing about their anxieties and receiving their advice really helped.

My experience isn't especially unusual. The PhD experience harbours many risk factors for mental ill health, including feeling lonely and isolated, and doubting your own abilities. PhD students face regular academic criticism and encounter unexpected challenges: experiments don't work; and ethics applications and papers get rejected. Often, universities provide no dedicated, preventive mental-health support for PhD students. We struggle in silence, and our engagement in our research deteriorates. Some take a leave of absence or drop out.

NATURE.COM
How one scientist cycled her way through a PhD:
go.nature.com/2yalcxy

COMMUNITY SPIRIT

In June 2018 I was still struggling, but I entered the university's Three Minute Thesis competition because my supervisors encouraged me to, and I was eager to please them. The competition challenges students to communicate their research topic and its significance to a non-academic audience in three minutes. ▶

I reached the final, and managed to keep going despite awful microphone feedback. My fellow finalists were all very supportive of each other, and my friends and supervisors came to watch. For the first time, I felt part of a vibrant and supportive postgraduate research community and had confidence in my abilities. I was able to make decisions that advanced my research project.

Finding support from others who were going through the PhD experience in my department and elsewhere — in other words, finding a community — was what helped me most when I struggled with mental ill health. I decided to try to create the same for others.

I thought about what I would have told myself six months previously that might have helped. When I asked other students the same thing, I discovered that everyone had found something challenging and had a corresponding message and advice. I sounded out the other Three Minute Thesis finalists, and some friends, to see whether they thought a workshop to share this information would be valuable and whether they would contribute. They responded positively, and one also suggested producing written resources, such as a booklet or online resource. As a result, I founded the 'How to survive your PhD (and enjoy it)' initiative as a way for postgraduate students and early-career researchers at the University of York to share their knowledge and experience with those at earlier stages of their PhDs.

"Finding a community was what helped me most when I struggled with mental ill health."

Its focus was information-sharing rather than mental health, but I hoped that it would offer a further benefit to students who might be struggling, by helping them to feel connected and highlighting opportunities for support.

GUIDE FOR SURVIVAL

I pitched the project to the university's researcher-development team and graduate research school, and to dedicated postgraduate colleges and student associations. I recruited volunteers across various disciplines, and from different faculties and departments; some were part-time students, some full-time. They represented all stages of the PhD journey, from first year to final year, and the first year of a postdoctoral fellowship. We identified common themes in the experience and developed a guide, web pages (see www.york.ac.uk/survive-your-phd) and two workshops that are open to all University of York PhD students through the university's researcher-development platform.

The guide and additional web content were written between June and August 2018, with contributions from 25 students and postdocs. I developed the content in collaboration with a committee of 10, whose energy and help made the project possible. To ensure a single voice, the text was then written up by a PhD student with journalistic experience. Two others produced layouts and images, and the committee proofread text and agreed the final layout.

University bodies signed off on the guide and web content and funded the printing, but they did not influence the content. They put me in touch with the university's web-content producer, who developed the web pages on the graduate-school website for us.

We held the first workshop in October 2018, and the second in February 2019. The workshops were largely the same, although the second one included some extra content about the international student experience, and about managing caring responsibilities. We used 10 presenters in each, and 81 students attended, of whom 98% found the session useful and 85% felt more confident afterwards about surviving and enjoying their PhDs. Like me, the participants particularly appreciated the opportunity to talk about their anxieties, finding peer support reassuring.

We are now also developing 'survival' workshops on the viva, fieldwork and international PhD-student experiences, and we're creating a sign-up system so that students can register their interest as workshop speakers and coordinators. This will ensure rolling recruitment, so that the project always has enough people involved in it from the various year groups and departments.

WHAT YOU SHOULD KNOW

Here's what I want current and future PhD students, and universities, to take away from my experiences:

- Students and universities should be candid about the challenges as well as the benefits of doing a PhD.
- Students' knowledge of the PhD process can and should be shared with those who are starting their PhDs — and not just in their own departmental silo. University backing is needed to help get peer-support initiatives off the ground and keep them going.
- Students can go to their doctoral training and support services and ask them what support is available for mental health and well-being. They can ask for help developing peer-support workshops across the university (not only in a single department) and promoting activities to students.
- Universities should work with PhD students to provide environments that reduce the risk factors for mental ill health, that help students to recognize when their mental health is being adversely affected, and that put them at ease about asking for help. ■

Sarah Masefield *is an occupational therapist and has expertise in patient and public-involvement research. She is currently writing up her PhD research on the use of primary health care by mothers with preschool disabled children.*

I AM NOT THE HIVE MIND OF TRANSETTI PRIME

A difficult decision.

BY STEVEN FISCHER

ILLUSTRATION BY JACEY

I am not a young girl staring up at the stars, grasping her father's hand and holding it tight.

She looks at his face, exhausted but smiling, then back again at the blanket of ink stretched above. She loves the stars on this side of the station, even though she rarely gets to see them. Even though a day off means Father will have to work extra shifts for the next week and a half.

But today that doesn't matter, because he's with her looking at the stars. At Sirius and Ursa Major, and even the pale white glimmer of Sol. She grins because these are their stars, here beneath the hum of the station's engines, purring like the tabby cat that stalks their crowded hab onboard.

As the station spins, the stars swirl slowly, tracing eon-worn ruts across the dark sky. She knows them all, or at least all with names, just as her father and his father before him.

Today she smiles, though, because she has found a new star, flickering dimly over the metal swell of the station's massive grav drive. She's waited three cycles to tell Father, because she wanted to be certain, rehearsing the named stars over and over in her mind. But this one is not named, nor does it belong to the myriad lists of those marked by just numbers.

It's a new star. Their star. Just like all the others.

Her smile widens as their star splits into two.

I am not the technician daydreaming in his chair when the cluster of red appears on the long-range sensors.

The alarms wake him quickly, not loud, but persistent. A steady, slow beep growing in volume, at first in rhythm with, then falling far behind, the beating of his heart.

It's probably a drill or a glitch in the system. Maybe debris from a wreck or a derelict vessel. He queries the endless transit databases, but there are no registered ships in that sector, and the pattern on the screen is too regular and moving too fast.

More than one signal, so not a wrecked vessel, too clear and too crisp for a cloud of debris.

There are few things in space that move that quickly. Even fewer that would trigger the system from this distance.

He sighs, and his hand goes to his neck, fingers threading across the port beneath his skull base and the smaller, blue gem set in his skin just below. His eyes close as he thinks of the matching stone in the skin of his lover half a system away. He whispers a silent prayer of thanks that their transport isn't returning today, then he forces the system to run the numbers again.

There is no glitch, but he knew that already.

I am not the admiral roaming the bridge, inspecting data displays and bulkhead stations nestled deep within the station's metal heart.

She turns her head at the sound of hurried footsteps, a young ensign racing across the gangplank, carrying a terminal in his hands.

"Easy, soldier," she whispers, taking the display from him, a star chart with a cloud of crimson dots hovering above.

Her eyes harden into a stone-heavy grey as they sweep across the tangles of projected trajectories, and she feels the station's core come alive around her.

There are protocols for these things, put in place by intelligences far greater than hers, and the station has already made up its mind. Somewhere in the tangles of steel that surround her, thousands of missiles are dropping into place, preparing to fire the moment the station determines it has no other options.

But she doesn't need the computational power of 2 million human minds to confirm that for her. There are no other options, and their course has already been set.

She stares at the ensign and the emblem on his chest — a ring of stars inside a shield of iron — and thinks of another young officer aboard another station, bearing not just her same emblem, but her same name as well. Because the shield that guards that ring of stars is not composed just of metal and missiles, but of breath, and blood, and soldiers with sons and mothers of their own.

I am not the hive mind of Transetti Prime. Two million voices compiled into one perfect thought. Four million eyes blending into one omniscient vision.

I may have been that only seconds ago, when I triggered alarms and opened my hangars, and set every ship in its berth loose with all they could carry.

I may have been that when I locked all my doors, keeping those inside safe from each other, at least, as I could not keep them safe from what waited outside.

I may have been that when I dropped a thousand warheads into silos, preparing to return the salvo headed my way, if in the end I could do nothing to stop it.

But I am not that anymore.

I am a grizzled veteran standing at her post, unable to escape the horrors of war, but afraid only because she knows her child will face them, too.

I am a sensor technician sitting in a chair, sending a final message to his lover away for work, grateful that only one of them will die today.

I am a child, standing on the bow of a station, arms wrapped around a father who's crying for reasons she can't understand, awestruck by their good fortune and the cloud of beautiful stars streaking down around them.

I am two million lives that are already dead, and the millions more that will die if I follow my directives.

I close the ports along my bow and rest my missiles still and silent in their tubes as the first of the warheads detonate against my hull, light like starfire consuming all around me. ■

Steven Fischer *is a resident physician in the Pacific Northwest of the United States. You can read more of his work at www.stevenbfischer.com or follow him on Twitter @stevenfischersf.*

➲ NATURE.COM
Follow Futures:
@NatureFutures
go.nature.com/mtoodm